University of Cincinnati

Date: 8/12/2015

I, Samuel Vaughn , hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Immunology.

It is entitled:

PXK and Lupus: Novel Immunobiology for a Lupus-Risk Gene

Student's name: **Samuel Vaughn**

This work and its defense approved by:

Committee chair: John Harley, M.D., Ph.D.

Committee member: Fred Finkelman, M.D.

Committee member: David Hildeman, Ph.D.

Committee member: William Ridgway, Ph.D.

Committee member: Stephanie Ware, M.D., Ph.D.

19626

PXK and Lupus: Novel Immunobiology for a Lupus-Risk Gene

Dissertation

Presented in Partial Fulfillment of the Requirements for the Doctor of Philosophy in The Graduate School of The University of Cincinnati College of Medicine

By

Samuel Edward Vaughn, B.S.

Immunology Graduate Training Program

The University of Cincinnati

2015

Dissertation Committee:
John Harley, Advisor
David Hildeman
William Ridgway
Fred Finkelman
Stephanie Ware

Abstract

Systemic lupus erythematosus (SLE) is an autoimmune disease characterized by multi-organ damage. There is a strong evidence for genetics playing an important role in the development of disease. Genetic association studies have begun to identify many SLE-associated genetic loci, yet for very few of these loci is the specific pathogenic mechanism of disease described. We undertook a genetic fine-mapping study of the *PXK* locus, which our group had previously identified as being associated with SLE. In the course of performing the fine-mapping study, we also developed a streamlined analytic approach. By using existing genetic analysis software suites, linking them with small Linux scripts and parallel processing multiple statistical models simultaneously, we were able to greatly facilitate our genetic analysis. For the PXK locus, we identify a 257kb haplotype associated with SLE. The strongest association was found at rs6445972 with $P < 4.62 \times 10^{-10}$, OR 0.81 (0.75–0.86). We confirmed this finding with Bayesian analysis, confirming a 95% credible set consisting of 172 variants. We found one association at the locus. In addition to the genetic fine-mapping results, we also demonstrate that PXK colocalizes with the B cell antigen receptor (BCR) following BCR-crosslinking. shRNA-mediated knockdown of PXK resulted in delayed BCR internalization. Cell lines derived from patients carrying the SLE-associated risk variant displayed a delay in BCR internalization. Given the role for BCR internalization in B cell fate determination and survival, our results support new candidate mechanisms for SLE-disease risk, including BCR trafficking and BCR-mediated B cell activation.

Acknowledgements

I would like to dedicate this dissertation to my wife, Rachel, who has supported me through this entire process. I would like to thank my children, Alessia, Matthew and Dylan, who bring me daily joy. I would like to thank the rest of my family, especially my parents, John and Marla Vaughn. They have been guiding lights in my life and have always been my biggest cheerleaders. I would also like to thank Isaac Harley, who inspired this project and helped me endlessly from beginning to end. I thank my advisor and the members of my committee for their guidance and suggestions. They have made this a much better work. I would like to thank all the members of the Harley lab who have been good colleagues, mentors, and friend. Finally, I owe a special thanks to Leah Kottyan. Leah has been a patient mentor, a passionate scientist, and a wonderful friend.

Table of Contents

List of Tables and Figures

Abbreviations

Anti-phospholipid (aPL)

Antigen presenting cell (APC)

B cell antigen receptor (BCR)

Bone marrow (BM)

Epstein-Barr virus (EBV)

Fas-mediated programed cell death (PCD)

Genome-Wide Association (GWA)

Immunoreceptor tyrosine-based activation motifs (ITAMs)

Immunoreceptor tyrosine-based inhibitory motifs (ITIMs)

Interferon (IFN)

Large Lupus Association Studies 1 and 2 (LLAS-1 and LLAS-2)

Lymphoblastoid Cell Line (LCL)

Major Histocompatibility Complex (MHC)

MHC class II-containing compartment (MIIC)

Micro-RNA (miRNA)

Minor allele frequency (MAF)

Protein tyrosine kinases (PTKs)

Protein tyrosine phosphatases (PTPs)

Single Nucleotide Polymorphism (SNP)

Spleen Tyrosine Kinase (SYK)

Src homology 2 (SH2)

Systemic lupus erythematosus (SLE or lupus)

T lymphocyte receptor (TCR)

Toll-like receptor (TLR)

Type 1 diabetes (T1D)

Selected Genetic Definitions

Minor allele frequency (MAF)- The frequency at which the less common allele occurs at a specific genomic position.

Hardy-Weinberg equilibrium (HWE)- Simply stated, allele frequencies in a population should remain constant from generation to generation in the absence of outside pressures. Deviations from HWE can be estimated within a genetic dataset, identifying potential bias within the data.

Bayes Factor (BF)- Used in Bayesian statistics, it is a measure of the relative evidence in the data.

Posterior probability- A measure of the probability in Bayesian statistics of the conditional probability of a random event after accounting for background information.

Single nucleotide polymorphism (SNP)- A DNA sequence variation at a single genomic position that varies within the population as a whole.

Linkage disequilibrium (LD)- The non-random association of alleles at a genetic locus. LD is a measure of the probability that a given allele will be inherited in tandem with another specified allele.

Imputation- A statistical genetic technique used to fill in missing genotyping data by comparing the genotype of a specific individual with the genotypes of a control cohort. Using the frequency of known genotypes of the control group permits a “best guess” approach to filling the missing genotypes of the specific individual.

Genome wide association study (GWAS)- A genetic association study using genotype information from generally a large, unbiased selection of genotyped SNPs across the entire genome.

Conditional analysis- A method to identify the specific influence of a given allele on the dataset as a whole, by including the genotype at that allele as a covariate in the overall analysis.

Chapter 1: Background and Introduction

The overarching theme of the immune system is balance. Our immune system recognizes and reacts to potentially harmful stimuli, whether a pathogenic virus, such as the flu, or a mutated cell that is no longer identified as being a part of our “self”. At the same time, it also recognizes but tolerates the food and myriad of commensal organisms that colonize our bodies. After recognition of something dangerous, a finely balanced series of signals leads to neutralization of the threat. The virus is killed, as are those cells that harbor it. The mutated cell is destroyed, averting the potential cancer. Importantly, following the elimination of the specific threat, the immune system is restored to its basal level. This restoration occurs through a finely tuned series of steps. This balance, this homeostasis, is the key to our survival and maintenance of health.

Fundamentally, autoimmunity occurs as a consequence of the loss of homeostasis. Sometimes this occurs within an isolated organ system, such as the thyroid as in Hashimoto’s thyroiditis- the immune system targets the cells of the thyroid, leading to loss of thyroid function and subsequently, disease. This loss of balance can also occur on a system wide level, as occurs in Systemic Lupus Erythematous (SLE) or lupus, in which widespread immune activation leads to multiple organ damage and chronic disease. In this introduction, I will describe and compare a few of the various forms of autoimmunity, specifically focusing on lupus.

General principles of autoimmunity

There is both direct and indirect evidence for autoimmunity. Throughout this dissertation, autoimmunity is defined as an immune response directed against self. This definition is important because we consider autoimmunity as a primary driver of disease. Strong evidence that autoimmunity is a primary disease process and not an epiphenomenon is the transfer of antibody resulting in disease. A prime example of this is maternal-fetal transfer of antibody where the transfer of a mother's autoantibodies to her fetus results in disease, such as in Grave's disease (GD) [1], myasthenia gravis (MG) [2], lupus (leading to both dermatologic and cardiac disease in the infant)[3] and Sjogren's (SS)[4]. Direct disruptions of healthy processes mediated by autoantibodies can also be observed in vitro as demonstrated by disruptions of the coagulation cascade. Both cold agglutinins and antiphospholipid (aPL) antibodies are capable of this disruption. Cold agglutinins, autoantibodies directed against polysaccharide complexes found on the surface of red blood cells, can be isolated from a patient's blood and hemolysis measured in vitro. aPL antibodies are directed against blood proteins involved in coagulation [5, 6]. The presence of these antibodies is predictive of an increase of thrombotic events. The aPL antibodies can be detected in vitro by a prolongation of the clotting times in coagulations assays.

Indirect evidence of the autoimmune response as a primary mediator of disease is found in animal models of human autoimmune diseases; either through recapitulation in animal models following immunization with suspected causal antigen, or through naturally occurring disease. Examples of the former include autoimmune thyroiditis in mice [7, 8] and the experimental autoimmune encephalomyelitis (EAE) mouse model of

multiple sclerosis (MS) [9]. Naturally occurring disease in mice include those with lupus-like phenotypes (NZB/NZW [10, 11], MRL/lpr [12], BXSB [13]) and the diabetic NOD mouse [14].

Finally, circumstantial evidence supports the autoimmune disease process as well. The presence of autoantibodies is an important aspect of the differential diagnosis, but natural autoantibodies are common. Family history can be an important factor as autoimmunity does run in families. Autoimmune disorders also demonstrate clear gender bias, with autoimmunity overall occurring more often in females than males. Within the spectrum of individual autoimmune diseases however, there is more variation which can be useful in diagnosis, as SS occurs only 5% of the time in men, while ulcerative colitis (UC) and diabetes mellitus (DM) have a fairly even male to female distribution [15]. The autoimmune reaction as a primary mediator of disease is thus supported by multiple pieces of evidence.

Autoantibodies in autoimmunity

As discussed above, autoantibodies are a common feature of some autoimmune diseases and are often a first step in the diagnosis. Autoantibodies cause disease by multiple mechanisms. Broadly, they can act in a similar capacity to regular antibodies, as facilitators of antigen processing and presentation by antigen presenting cells (APCs) via opsonization, or to activate the complement cascade. Autoantibody binding and subsequent opsonization are important steps in the phagocyte-dependent presentation of nuclear antigens [16] and in the maintenance of the autoimmune response in immune

thrombocytopenic purpura (ITP) [17]. Immune complex binding by the complement protein C1q is the first step of the classical complement activation pathway. As autoantibodies also form immune complexes (ICs), it is no surprise that they also can activate the complement cascade. Complement binding to ICs links the innate and adaptive immune responses. The presence of complement during antigen presentation to B cells lowers the activation threshold [18].

For most autoimmune diseases, even though there is strong evidence that autoantibodies are important in the pathogenesis of disease, actual demonstration of autoantibodies having a causal role in disease pathogenesis is rare. Anti-glomerular basement membrane (anti-GBM) disease is one of the few exceptions. The Goodpasture antigen is directed against type IV collagen. Antibody binding leads to severe GBM damage and involves complement activation as demonstrated by C3 deposition [19, 20].

In addition to the traditional functions of antibodies, autoantibodies directed against receptors can also have a functional impact on receptor signaling. In GD, autoantibodies target the TSH receptor resulting in receptor activation and subsequent hyperthyroidism [21]. Myasthenia gravis (MG) is an autoimmune disease marked by autoimmune mediated disruption of the neuromuscular junction that results in increasing muscle weakness and fatigability. In MG, autoantibodies both block receptor signaling and fix complement, leading to receptor degradation. Binding of autoantibodies to the acetylcholine receptor (AChR) results in receptor blockade [22]. As the disease progresses, a significant reduction of AChR from the neuromuscular junction is

observed as a result of complement mediated destruction. Autoantibodies can thus have both direct and indirect involvement in autoimmune disease pathogenesis and are not present solely as an incidental byproduct of an autoimmune inflammatory reaction.

Systemic Lupus Erythematosus: Clinical findings

A diagnosis of lupus requires a patient to meet 4 of 11 criteria as defined by the American College of Rheumatology (ACR) [23, 24], see Table 1. The wide range of possible symptoms underscores the heterogeneous nature of lupus. These criteria do not necessarily need to occur simultaneously and often symptoms wax and wane over the course of the disease. The main symptoms of lupus as experienced by the patient can be divided into two groups- constitutional symptoms and those that relate to organ specific involvement. The frequencies of various manifestations are outlined in Table 2. The main constitutional symptoms are fatigue, myalgia, fever and weight loss. All are very common in lupus, with estimates ranging from 50-100% of patients experiencing these constitutional symptoms. While constitutional symptoms are very common in lupus, they are also very non-specific.

Table 1: ACR criteria for the clinical diagnosis of SLE.

Criterion	Definition
1. Malar rash	Fixed erythema, flat or raised, over the malar eminences, tending to spare the nasolabial folds
2. Discoid rash	Erythematous raised patches with adherent keratotic scaling and follicular plugging; atrophic scarring may occur in older lesions
3. Photosensitivity	Skin rash as a result of unusual reaction to sunlight, by patient history or physician observation
4. Oral ulcers	Oral or nasopharyngeal ulceration, usually painless, observed by physician
5. Arthritis	Nonerosive arthritis involving 2 or more peripheral joints, characterized by tenderness, swelling, or effusion
6. Serositis	a) Pleuritis--convincing history of pleuritic pain or rubbing heard by a physician or evidence of pleural effusion OR b) Pericarditis--documented by ECG or rub or evidence of pericardial effusion
7. Renal disorder	a) Persistent proteinuria greater than 0.5 grams per day or grater than 3+ if quantitation not performed OR b) Cellular casts--may be red cell, hemoglobin, granular, tubular, or mixed
8. Neurologic disorder	a) Seizures--in the absence of offending drugs or known metabolic derangements; e.g., uremia, ketoacidosis, or electrolyte imbalance OR

	b) Psychosis--in the absence of offending drugs or known metabolic derangements, e.g., uremia, ketoacidosis, or electrolyte imbalance
9. Hematologic disorder	a) Hemolytic anemia--with reticulocytosis OR b) Leukopenia--less than 4,000/mm3 total on 2 or more occasions OR c) Lyphopenia--less than 1,500/mm3 on 2 or more occasions OR d) Thrombocytopenia--less than 100,000/mm3 in the absence of offending drugs
10. Immunologic disorder	a) Positive LE cell preparation OR b) Anti-DNA: antibody to native DNA in abnormal titer OR c) Anti-Sm: presence of antibody to Sm nuclear antigen OR d) False positive serologic test for syphilis known to be positive for at least 6 months and confirmed by Treponema pallidum immobilization or fluorescent treponemal antibody absorption test
11. Antinuclear antibody	An abnormal titer of ANA by immunofluorescence or an equivalent assay at any point in time and in the absence of drugs known to be associated with "drug-induced lupus" syndrome

4 of 11 criteria are required, either sequentially or simultaneously, for a clinical diagnosis of SLE. Table adapted from Tan et al., Arthritis Rheum. 1982 [23], and Hochberg, Arthritis Rheum,1997 [24].

Table 2: Frequency of selected manifestations of lupus

Manifestation	Frequency (%)
Constitutional symptoms (fatigue, fever, weight loss)	90-95
Mucocutaneous involvement (malar rash, alopecia, mucosal ulcers, discoid lesions, etc.)	80-90
Musculoskeletal involvement (arthritis/arthralgia, avascular necrosis, myositis, etc.)	80-90
Serositis (pleuritis, pericarditis, peritonitis)	50-70
Glomerulonephritis	40-60
Neuropsychiatric involvement (cognitive impairment, depression, psychosis, seizures, stroke, demyelinating syndromes, peripheral neuropathy, etc.)	40-60
Autoimmune cytopenia (anemia, thrombocytopenia)	20-30

The varied manifestations of lupus and their respective frequencies underscore the heterogeneous aspect of this disease. Table adapted from Dall'Era et al., Kelley's Textbook of Rheumatology, 2013 [317].

Organ specific manifestations of lupus can also be shared amongst many diseases, especially other autoimmune diseases. This underscores the importance of the ACR criteria, as it is the unique combination of organ system involvement and laboratory findings that support the diagnosis of lupus. Most organ systems can be involved in some way or another in lupus and will be reviewed briefly below.

Joint symptoms including inflammatory arthritis are very common and often part of the clinical picture at initial presentation. Lupus arthritis is rarely deforming, in contrast with rheumatoid arthritis (RA), which often leads to sever joint erosion and deformation [25, 26].

Another of the most common clinical findings at initial presentation is involvement of the skin. The most common rash of lupus is the butterfly rash, marked by erythema on the cheeks and nose sparing the nasolabial folds. More severe discoid lesions can also occur and lead to scaring. Alopecia is also common [27]. Oral and nasal

mucocutaneous lesions also occur and tend to be less painful than those associated with Herpes simplex [28]. The Raynaud phenomenon, blanching or cyanosis of distal skin as a result of vasoconstriction in response to cold temperature, particularly in the digits of the hands and feet, is also common. Again, the Raynaud phenomenon is not lupus specific as it can occur both primarily and secondarily, especially in relation to other connective tissue diseases and scleroderma.

Kidney involvement is very common, with 40-60% estimates of lifetime occurrence in lupus patients, with renal disease being one of the major factors in the long-term morbidity and mortality in SLE. Immune complex-mediated glomerulonephritis is the most common form of kidney involvement [29]. Due to the high prevalence of kidney disease and the potential for severe complications, close monitoring of renal health is recommended [14, 30]. The presence of kidney disease is usually indicative of a poorer outcome [31].

The gastrointestinal tract can be involved both primarily from lupus itself, or more commonly, secondarily due to medication side effects. Non-steroidal anti-inflammatory drugs (NSAIDs) and glucocorticoids, both commonly used in lupus, have significant gastrointestinal (GI) side effects, including dyspepsia and peptic ulcers. Dysphagia is the most common GI complaint in lupus. Decreased peristalsis in the esophagus is the most frequent cause and can occur in both SLE and scleroderma, though SLE rarely affects the lower esophageal sphincter, which is common in scleroderma [32].

The most common pulmonary symptom is pleuritis but can also include pleural effusions, chronic interstitial lung disease, pulmonary arterial hypertension and others.

aPL antibodies are associated with thromboembolic lesions in the lungs. Dyspnea is a common complaint with clinical findings of abnormal pulmonary function tests consistent with restrictive abnormalities. The most common cardiovascular manifestations of lupus include pericarditis and valvular abnormalities such as valvular thickening and regurgitation. In addition, patients with lupus are at increased risk for coronary artery disease. aPL antibodies are also associated with coronary artery thrombosis.

Neurologic involvement can range from neuropsychiatric symptoms such as depression to neurosensory disruption like peripheral neuropathy [33]. The precise pathogenesis is often unknown. Neurological disease in lupus can be roughly divided into vascular in origin or inflammatory in origin. aPL antibodies are associated with CNS thromboembolic events [34]. Inflammation can lead to increased permeability of the blood-brain barrier, cytokine production, etc. Clinical manifestations of neurologic involvement in lupus include cognitive dysfunction such as difficulty in thinking, memory and concentration [35, 36]. Additionally psychosis, depression and anxiety can all occur in patients with lupus [37, 38]. Specific nerve involvement can occur as in MS, including optic neuritis and myelitis. Autoantibodies suggestive of lupus plus other relevant history can help differentiate the two. Peripheral neuropathy can also occur [39]. The most common ophthalmologic complaint is keratoconjunctivitis sicca (KCS), which often occurs with secondary SS [40].

Hematologic disruptions are common in lupus, most commonly cytopenias including leukopenia, anemia and thrombocytopenia [41]. Anemia is most often anemia of chronic disease but autoimmune hemolytic anemia also occurs. Leukopenia can be

primary or secondary as a result of medication, especially glucocorticoids. Thrombophilia can occur most often in the context of aPL antibodies, though advanced kidney disease can also increase the risk of thromboembolic events [42]. Peripheral lymphadenopathy is another common finding [43]. Vasculitis can occur in SLE with presentation dependent upon localization of the lesion [44].

Autoantibodies in lupus

The prototypical immunological finding in lupus is the production of autoantibodies. Autoantibodies are a distinctive feature of many autoimmune diseases, however, and do not provide direct evidence of autoimmunity as they can be found both in healthy individuals [45] and prior to overt symptomology in lupus [46], RA, antiphospholipid syndrome, and type I diabetes mellitus, though in the later case they may be highly predictive of disease [47].

The most common autoantibodies in lupus are the anti-nuclear antibodies (Table 3). These are autoantibodies targeting nuclear components of the cell and are highly sensitive for lupus; over 95% of SLE patients are ANA positive. The antibodies are detected in a fairly unique manner for a clinical test, as HEp-2 cells, a human epithelial cell line, are treated with dilutions of serum from patients where a positive test is indicated by a nuclear distribution of immunostaining [48-50]. ANAs are not as specific, however, as ANAs can also be positive in RA, scleroderma, polymyositis, SS, and are also found in healthy controls [51, 52]. Following a positive ANA result, patients are usually screened for further autoantibodies. Anti-Sm is one of the most specific for

autoantibodies for lupus, with a specificity of 98% [53], with anti-dsDNA used most commonly for confirmation.

Table 3: Common autoantibodies in lupus.

Antigen Specificity	Prevelance (%)	ANA?	Main Clinical effects
Anti-double-stranded DNA	70-80	Yes	Kidney disease, skin disease
Nucleosomes	60-90	Yes	Kidney disease, skin disease
Ro	30-40	Yes	Skin disease, kidney disease, fetal heart problems
La	15-20	Yes	Fetal heart problems
Smith (Sm)	10-30	Yes	Kidney disease
NMDA receptor	33-50	No	Brain disease
Phospholipids	20-30	No	Thrombosis, pregnancy loss
α-actinin	20	No	Kidney disease
C1q	40-50	No	Kidney disease

Chart of the most frequent autoantibodies in lupus detailing their prevalence, whether or not they are antinuclear, and the main clinical effects with which they are associated. Adapted from Rahman and Isenberg, NEJM, 2008 [54]. References: dsDNA: [55-57], Nucleosomes: [58], Ro: [59], [60], La: [59], Sm [61], NMDA receptor: [62], [63], phospholipids: [64], α-actinin: [65], [66], C1q: [67].

The generation of auto-reactive clones during lymphocyte development, in both B and T cell pools, is common. Upward of 50% of BCR/TCR rearrangements are responsive to self-antigen [68-70]. Given the high rate of autoreactive lymphocyte production, yet the low overall rate of autoimmune disease prevalence of less than 5% [71], it is apparent that the immune system is able to fine-tune the lymphocyte repertoire. Four general mechanisms for limiting the impact of autoreactive clones have been described [72]: 1) Eliminate cells with a self-reactive receptor. 2) Edit the self-

reactive receptor. 3) Alter the responsiveness of the self-reactive cell through gene expression or other cell-intrinsic changes. 4) Limit the availability of outside help (such as co-stimuli, growth factors, etc.), reducing the ability of self-reactive clones to activate in the periphery. The relevance of this model to understanding the pathogenesis of lupus is evidenced by the examples from both mouse models and human disease of disruption of each of these pathways leading to autoimmunity and in many cases either lupus or a lupus-like disease. The MRL-lpr, or Fas^{lpr} mice clearly demonstrate the consequences of the failure to eliminate self-reactive lymphocytes. These mice develop severe lymphoproliferation and a lupus-like phenotype due to a mutation in the *Fas* gene, preventing apoptotic elimination of self-reactive lymphocytes [73-75]. Homozygous *lpr* mice have a more severe phenotype than the background MRL mice, which still have a late autoimmune syndrome [76]. Crossing the *lpr* mutation onto congenic inbred strains resulted in autoantibody production, but in a less severe overall phenotype [76]. This suggests that *lpr/fas* contributes to autoimmune disease, but that the background genetic environment also significantly contributes to disease pathogenesis [77, 78]. While there are further examples of disruptions of elimination of self-reactive lymphocytes leading to autoimmunity, including *Bim* deficient mice developing autoantibodies [79], and mutations in *AIRE* leading to autoimmune polyendocrine syndrome (APECED), a severe autoimmune disorder [80, 81], there is little evidence for a failure in receptor editing leading to autoimmunity. Interestingly, T cells are primary mediators in APECED [82], but B cells play a critical role in regulating the T cell response [83] with limited support for rituximab as a potential treatment option

[84]. Thus it seems that most of the regulation of autoimmune clones falls on the cell-intrinsic and cell-extrinsic tuning mechanisms.

B cells as targets for lupus therapies

B cells have been a main focus of new lupus therapy development. Belimumab, an anti-BAFF therapy targeting B cell development was the first new drug approved by the FDA for the treatment of lupus in over 50 years [85]. Belimumab has been shown to reduce autoantibody titers, reduced plasma cell numbers without greatly effecting memory B cells or T cell populations [86]. There is evidence that belimumab was effective in treating lupus associated renal disease, although patients with advanced disease were excluded [87]. Further study is obviously needed. Results with rituximab, a chimeric monoclonal antibody that targets CD20 positive cells, have been mixed. While some earlier studies suggested that rituximab may be efficacious in lupus [88-94] and rituximab has been shown to reduce overall B cell numbers in patients with lupus, lupus nephritis remained unchanged [95]. The later finding may be a consequence of study design [96], as meta-analysis of existing data supports rituximab as an effective in lupus nephritis [97]. Another treatment that initially showed promised was an anti-CD154 antibody. Treatment with anti-CD154 resulted in a decrease in activated B cells in patients with lupus and also a reduction in serum autoantibodies, proteinuria and overall lupus severity. The trial was stopped though as there was also an increase in thrombotic events [98].

B cell development

B cells arise from hematopoietic progenitors in the bone marrow and fetal liver. The BCR is the fundamental distinguishing feature of the B cell and the development and diversification of the BCR lies at the heart of B cell development [99, 100]. The BCR consists structurally of heavy and light chains encoded by the V, D and J gene segments in the H chain locus and V and J gene segments in the L chain locus. The diverse repertoire of BCRs produced arises from the ordered rearrangement of these segments [101-105]. The heavy chain is assembled first during the pro-B cell stage. Following rearrangement of the H chain locus, the newly transcribed heavy chain is paired with a surrogate light chain to form the pre-BCR [106, 107]. This surface expression of this receptor marks the pre-B cell stage of B cell development. Underscoring the central role of BCR signaling to B cell development and survival, only pre-B cells with functional heavy chains (those capable of signaling when paired with the surrogate light chain) proceed with rearrangement of the L chain locus. Those that fail this checkpoint undergo additional rounds of H chain rearrangement until either a functional heavy chain is produced or until the cell is eliminated [108]. Following successful rearrangement of the L chain locus and functional heavy/light chain pairing the BCR+ immature B cell leaves the bone marrow and enters the periphery en route to the spleen or other peripheral lymphoid organs where further maturation occurs. During this period between emergence from the bone marrow and subsequent maturation, the immature B cells are known as transitional B cells [109].

Transitional B cells and lupus

The Ig locus gives rise to the various Ig classes, including IgM, IgD, IgG, IgA and IgE. During maturation in the bone marrow, surface-bound IgM is the first class to be expressed. (This stage of B cell development culminating in functional IgM expression is also known as the antigen independent stage). Transitional B cells continue to express IgM along with low levels of IgD. There are two main subsets of transitional B cells: T1 and T2. These subsets are found in both mice and humans and can be distinguished by CD21/CD24 expression amongst CD19+IgM+ cells [110]. In the context of lupus and BCR signaling, transitional B cell are interesting in that they appear to be uniquely sensitive to BCR crosslinking and may mark an important step in B cell selection. T1 cells appear to be more susceptible to BCR-induced apoptosis than T2 cells in vitro [111, 112], which may be due to T2-specific rescue from apoptosis [113].

Transitional B cells are more abundant in blood from patients with lupus [114, 115]. This difference was found to be independent of absolute numbers, suggesting bone marrow production was normal [115]. This increase is not associated with disease activity [109]. In fact, transitional B cells are also increased in immunocompromised individuals including those with X-linked immunoproliferative disorder (XLP) [116]. The exact role of transitional B cells in lupus requires further research, but provides further evidence of the disruption in homeostasis underlying lupus [117].

Genetic and environmental factors in lupus pathogenesis

Given the heterogeneous nature of lupus, it is not the breakdown or loss of a single mechanism or system that leads to disease, but instead is most likely a result of the interaction of multiple factors in concert that result in the breakdown of homeostasis and subsequent overt manifestations of disease. The exact pathophysiological cause of lupus is still unknown, but various contributing factors have been or are just now beginning to be identified and will be discussed below. These potential causal factors can be broken down into two broad categories, namely environmental and genetic.

Environmental factors for the purpose of this discussion constitute those agents that are external in nature and include chemical exposure and infectious agents. Multiple environmental factors have been investigated as having a role in lupus pathogenesis including smoking, therapeutic drugs, UV exposure, EBV (which is discussed in Chapter 2), and uranium exposure. Cigarette smoking has been associated with both the development of SLE [118] and disease activity [119]. A study in a Japanese population suggests that this relationship may be due to differential production of reactive oxygen species [120].

Various drugs have been found to induce a lupus-like disease in patients including procainamide, hydralazine, minocycline and TNF-alpha therapy. Overall, symptoms of drug-induced lupus are similar to SLE and include fever, myalgia, rash, arthralgia and arthritis. The exact pathogenesis of drug-induced lupus remains unknown, but the underlying pathology appears to be very similar to SLE as they share many of the same features. Serological abnormalities are common, with autoantibodies, especially anti-ANA, occurring frequently. Thrombocytopenia and elevated gamma

globulins are also common. Interestingly, the lupus-inducing drugs are highly diverse with few structural similarities. In addition, the drugs are largely inert; there is no evidence that a specific property of the chemical compound causes lupus. Finally, drug-induced lupus occurs weeks or months after treatment initiation, yet the lupus-inducing drugs reach steady state quite rapidly. These observations have lead to the hypothesis that altered drug metabolism is responsible for drug-induced lupus.

UV exposure is known to induce or exacerbate lupus skin lesions [121]. UV light acts on keratinocytes to secrete inflammatory cytokines including IL-1, IL-3, IL-6, GM-CSF and TNF-alpha [122, 123]. UV-induced apoptosis of keratinocytes causes autoantigens to cluster in blebs near the cell surface. These autoantigen-containing structures could presumably be a source for subsequent re-exposure leading to clinical flares [124].

Chemical and industrial exposures have also been found to increase the risk of lupus, including silica exposure [125] and uranium exposure [126]. Other chemical and industrial exposures such as mercury and hydrocarbons show some evidence of association and are the subject of further research [127].

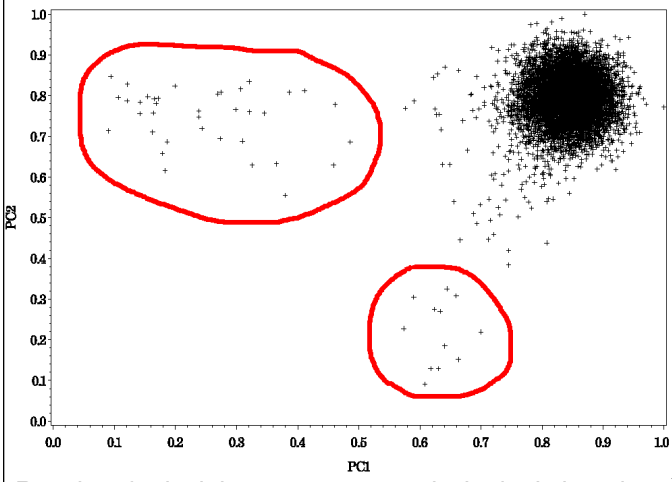
Statistical strategies to genetic association testing

Recent advances in genetic analysis have provided further support of the genetic underpinnings of SLE pathogenesis. This is reviewed more thoroughly in my published review from 2012 [128], included as chapter 2 of this dissertation. The following section will develop selected themes more specifically.

The underlying hypothesis of genetic association studies is that these associated variants serve as markers for actual causal variants. One of the basic ways to test this hypothesis statistically is with a chi-square test. Using the chi-square formula $X^2 = \sum \frac{(Observed\ value - Expected\ value)^2}{(Expected\ value)}$, we can calculate the difference between the observed and expected distribution of alleles at each variants. Under the null hypothesis, we expected the same allelic distribution at a given variant in both cases and controls. While this model allows us to quickly calculate a p value for the difference in allele distribution between cases and controls, is limited as we cannot introduce covariates into the model to account for some of the other potential differences between our groups, including ancestral makeup. While efforts are made to select populations as closely matched as possible, there is still the possibility that the case population has an underlying difference from the controls that will influence the chi-square distribution, confound the results.

One way to control for these differences in study populations is by calculating population substructure, or admixture, and adding this variable to the model. Admixture is the measure of the different ancestral population

Figure 1: Plot of first two principal components in individuals reporting European ancestry.



Results of principle component analysis depicting plot of first two principle components, revealing population substructure derived from underlying genetic similarities and differences. Individuals that have more genetic substructure in common are clustered together. Outliers represent individuals with significant differences from the rest of the group, and are circled in red.

makeup of an individual's genome [129, 130]. Two common ways by which this is done is through principal component analysis (PCA) and by direct admixture calculation. Principal component analysis calculates eigenvectors for each sample in the population [131]. These data can then be viewed on a scatterplot, giving a visual confirmation of underlying substructure. This analysis can be repeated multiple times, giving multiple principal components for each sample (Figure 1). These principal components are then taken into account in the final model, allow the model to control for whatever underlying structure exists. Similar to PCA, admixture can be directly calculated by using ancestral informative markers (AIMs), which have been validated to be able to distinguish between different ancestral populations. These AIMs are genotyped as part of the study, then admixture values are calculated for each ancestry of interest for each individual. In this manner, underlying admixture can be directly calculated for each sample in the study population.

As chi-square analysis does not permit the introduction of the admixture estimates in the association model, logistic regression is commonly used in association analysis. Logistic regression allows for the inclusion of additional covariates in the statistical model, including admixture estimates as described above. This permits a more accurate analysis, accounting for the potential error introduced by population stratification that may otherwise lead to false positives.

PXK

As the focus of my dissertation work is PXK, the following section briefly reviews the gene and protein. *PXK* encodes a PX domain containing serine/threonine kinase. It was first described in 2005 based on homology with other PX domain containing proteins [132]. The PX domain, or Phox-homology domain, was first identified in NADPH oxidase [133] and has been found to be a common motif especially in the sorting nexin (SNX) family of proteins [134, 135]. The PX domain interacts primarily with phosphoinositides, in particular phosphotidylinositol-3-phosphate (PI3P) [134], though the PX domain has also been shown to participate in protein-protein interactions as well [136]. PX domain proteins participate in a number of signaling and membrane-associated cellular activities including endosomal membrane sorting [137], controlling intercellular trafficking [138, 139], and trafficking of receptors into endosomes [140]. In addition to the PX domain, *PXK* also contains a region with high homology to a serine/threonine protein kinase domain. However, there are key highly conserved residues in ser/thr-protein kinases that are non-conserved in PXK. In addition, no kinase activity was detected when tested, suggesting that PXK is a dead kinase [141].

The function of PXK remains largely unexplored. Takeuchi et al provided the first experimental examination of the functional role of PXK [141]. They found that PXK localized to the endosomal membrane in COS-7 cells, and that site-directed mutagenesis of the PX domain or inhibition of PI3K with wortmannin completely eliminated this colocalization. Additionally, they showed that the presence of PXK accelerated the internalization and subsequent degradation of EGFR. PXK depletion

resulted in reduced internalization of the EGFR, most notably at the earliest time points, with a correction toward normal at later time points [141].

Our group first identified the PXK locus as a lupus-associated locus in 2008 [142]. Since that time, it has been confirmed as a lupus-associated locus [143], [144] and as a susceptibility locus in systemic sclerosis [145] and RA [146]. In each of these studies, the association was limited to individuals of Caucasian descent. Multiple studies have looked in particular at Asian populations, but have not found an association between overall SLE occurrence and the *PXK* locus [147, 148], though the former did identify an association between the *PXK* locus and autoantibody production [147]. While we were preparing our manuscript for publication, another group published a smaller fine mapping of study of the *PXK* locus [149]. This study confirms the genetic association with lupus in the locus, but focuses on another gene in the region, ABHD6, as being the causal gene. I will explore these findings and their conclusions further in both my published manuscript that follows, and in the discussion section at the conclusion of this dissertation. In summary, we found that there is a strong, single genetic association with lupus at the *PXK* locus. In addition, we found that PXK colocalizes with the BCR and that PXK knockdown resulted in delayed internalization of the BCR.

- *PXK* encodes a PX domain containing ser/thr kinase with critical non-conservation in the kinase domain.
- PXK is a member of the SNX family of proteins and has been found to regulate internalization of receptors.

- Variants in the *PXK* locus are associated with lupus, RA, and systemic sclerosis.

Of the specific lupus associated loci discussed above, my graduate work is focused on the *PXK* locus. I approached this locus from two directions simultaneously; from a genetic analysis standpoint, and from biological perspective. The next section is our review of the current state of lupus genetics. The following section is a presentation of the analytical strategies and tools we utilized to undertake a fine mapping of the *PXK* locus and other loci in our group as well. Our overall strategy was refined considerably over the past few years. I then present the published manuscript that presents the results of both our fine-mapping study and the biological studies focused on the impact of the SLE-associated genetic signal on PXK function within the B cell.

Chapter 2: Genetic Susceptibility to Lupus: the biological basis of genetic risk found in B cell signaling pathways

Introduction

Systemic lupus erythematosus (SLE or lupus) is an archetypical systemic autoantibody-mediated autoimmune disease. Autoantibody production in lupus leads to direct autoantibody-target tissue damage and to indirect autoantibody-autoantigen (immune) complex formation and deposition, both resulting in multisystem tissue injury. Lupus is highly variable both in severity and in the range of clinical manifestations, as evidenced by allowing classification as lupus with any combination of 4 of 11 criteria [23, 24].

Lupus is ten-fold more frequent in women and three- to seven-fold more frequent in non-European ancestries [150]. Despite its heterogeneity, SLE has a strong genetic component, with heritability estimates as high as 66%, a high sibling risk ratio (8-29), and a concordance rate in monozygotic twins (~30%) ten times higher than that of dizygotic twins (~3%) [151]. The most powerful evidence for a genetic component for lupus susceptibility, however, is the approximately 50 genetic variants that are now convincingly established to be associated with lupus (Table 4). In aggregate, these independent associations account for a minor fraction of the genetic risk of lupus [151], suggesting that other risk factors and better models of genetic effect remain to be discovered.

Table 4: Genes Associated with Lupus

Gene	Location	Odds Ratio	Best p-value	population	References	SNP with best p value (Risk allele) Ref	MAF of Best SNP	Causal variant (minor allele)
PTPN22	1p13.2	1.35	3.4×10^{-12}	EU, HA	[152-155]	rs2476601 (A) [154]	0.10	rs2476601 (A) PTPN22 R620W
FCGR2A, FCGR3B	1q23	1.59	9.1×10^{-7}*	EU, AA, AS	[153, 156-160]	rs1801274 (T) [153]	0.43	rs1801274 (T) FCGR2A H166R
NCF2	1q25	1.19	1.01×10^{-54}	EU, AS,AA, HA	[154, 161-165]	rs17849502 (G) [165]	0.12	rs17849502 (G) NCF2 H389Q rs17849501 (T) NCF2 A202A
CRP	1q23.2	0.49	9.2×10^{-14}	EU, AA	[166]	rs3093061 (G) [166]	0.19	
TNFSF4	1q25	1.46	2.5×10^{-32}	EU, AS, HA	[153, 154, 162, 167-169]	rs2205960 (**T**) [168]	0.35	
IL10	1q31-q32	1.19	4.0×10^{-8}	EU, AA, AS	[154, 170-172]	rs3024505 (A) [154]	0.16	Elk-1 preferentially binds lupus-risk allele rs3122605 (G)
C1q *C1r, C1s* *C2, C4*	1p36 12p13 6p21.3		Convincing Family Studies*	EU	[173-177]			Multiple, usually gene product deficiencies
TET3	2p13.1	0.75	6.9×10^{-17}	AS	[178]	rs6705628 (C) [178]	0.05	
RASGRP3	2p25.1-24.1	0.7	1.3×10^{-15}	AS	[168]	rs13385731 (C) [168]	0.13	
IFIH1	2q21135 12204	1.11	1.6×10^{-8}	EU, AA	[164, 179, 180]	rs13023380 (G) [180]	0.75	rs10930046 *IFIH1* R460H rs1990760 *IFIH1* A946T rs13023380 (A) reduces the affinity efficiency of nuclear extract binding
STAT4/STAT1	2q32.2-q32.3	1.55	5.17×10^{-42}	EU, AA, AS, HA	[153, 162, 167, 181-187]	rs7582694 (C) [168]	0.42	
CD80	3q13.3-q21	0.79	2.5×10^{-16}	AS	[168]	rs6804441 (A) [168]	0.31	
PXK	3p14.3	1.25	7.1×10^{-9}	EU	[153, 162, 169, 185, 188]	rs6445975 (C) [153]	0.32	
TREX1	3p21.31	44.65	8.5×10^{-11}	EU	[189, 190]	rs3135945 (A) [189]	0.01	TREX1 R11H and others (all uncommon or rare)
TMEM39A	3q13.33	0.72	8.62×10^{-9}	EU, AS	[191]	rs1132200 (A) [191]	0.01	
AFF1	4q21	0.81	8.3×10^{-9}	AS	[186]	rs340630 (A) [186]	0.44	
BANK1	4q24	1.3	2.642×10^{-13}	EU, AA, AS, HA	[162, 167, 192, 193]	rs10516487 (A)	0.13	rs10516487 (A) BANK1 R61H

						[193]		
IL2/IL21	4q26-q27	0.86	2.2×10^{-8}	EU, AA, AS	[161, 194, 195]	rs907715 (G) [194]	0.31	
TNIP1	5q32-q33.1	1.27	3.8×10^{-13}	EU, AA, AS	[154, 162, 168, 196-198]	rs7708392 (G) [153]	0.45	
PTTG1	5q35.1	1.22	1×10^{-10}	EU	[153, 199]	rs2431697 (C) [153]	0.38	
MIR146a	5q34	1.29	2.74×20^{-8}	AA, AS, EU	[199-201]	rs57095329 (G) [200]	0.21	rs57095329 affects binding of Ets1 to the promoter
HLA and other genes	6p21.3	2.35	1.27×10^{-51}	EU, AS, AA, HA	[153-155, 162, 168, 186, 196, 197, 200, 202-205]	rs1270942 (G[153]	0.2	
ATG5	6q21	1.25	5.2×10^{-12}	EU, AS	[153, 154, 162, 168]	rs548234 (G) [168]	0.33	
UHRF1BP1	6p21	1.49	44×10^{-9}	EU, AS	[154, 206]	rs13205210 (C) [206]	0.18	rs11755393 UHRF1BP1 R454Q rs13205210 UHRF1BP1 M1098T
TNFAIP3	6q23	1.72	1.3×10^{-17}	EU, AA, AS	[154, 162, 168, 186, 207-210]	rs2230926 (C) [168]	0.07	TT>A polymorphic dinucleotide; decreased NFκB binding by inhibiting DNA looping facilitated by NF-kB and SATB1
HIP1	7q11.23	1.43	1.3×10^{-8}	AS	[186]	rs6964720 (A) [186]	0.25	
IKZF1	7p12.2	0.72	2.8×10^{-23}	AS, EU	[162, 164, 168, 198]	rs4917014 (A) [168]	0.25	
JAZF1	7p15.2-p15.1	1.19	1.5×10^{-9}	EU	[154, 162]	rs849142 (T) [154]	0.49	
IRF5/TNPO3	7q32	1.54	3.611×10^{-19}	EU, AA, AS, HA	[153-155, 162, 211-214]	rs12537284 (A) [153]	0.19	Independent variants reduce IRF5 activity
XKR6	8p23.1	0.81	2.5×10^{-11}	EU	[153, 155]	rs6985109 (G) [153]	0.49	
BLK	8p23-p22	0.69	2.1×10^{-24}	EU, AA, AS, HA	[153, 168, 186, 215-217]	rs7812879 (A) [168]	0.12	
LYN	8q13	0.77	5.4×10^{-9}	EU, AA, AS	[153, 218]	rs7829816 (C) [153]	0.18	
LRRC18, WDFY4	10q11.23	1.24	7.2×10^{-12}	AS,EU	[162, 168, 219, 220]	rs1913517 (A) [168]	0.33	rs877819 disrupts YY1 binding and decreases expression of WDFY4
ARID5B	10q21.2	0.85	5.1×10^{-11}	AS	[168]	rs4948496 (C) [168]	0.15	
CD44,PDHX	11p13	0.71	4.0×10^{-12}	EU, AS,	[169, 221]	rs507230	0.44	

				AA		(G) [169]		
PHRF1/IRF7/KIAA1542	11p15.5	0.78	$3x10^{-10}$	EU, AA	[153, 154, 162, 168]	rs4963128 (G) [168]	0.07	IRF7 Q412R increases IRE transcription
DDX6, TREH	*11q23.3*	1.29	$1x10^{-16}$	AS	[222]	rs4639966 (G) [222]	0.33	
ETS1	11q23.3	1.37	$1.8x10^{-25}$	AS	[168, 186, 196, 198, 219]	rs6590330 (A) [168]	0.41	
GPR19, CDKN1B	12p13.1-p12	0.79	$3.8x10^{-17}$	AS	[168]	rs10845606 (C) [168]	0.27	
DRAM1	12q23.2	0.84	$9.4x10^{-12}$	AS	[168]	rs4622329 (A) [168]	0.34	
SLC15A4	12q24.32	1.26	$1.77x10^{-11}$	AS	[168, 198]	rs1385374 (A) [168]	0.25	
ELF1	13q13	1.26	$1.5x10^{-8}$	AS	[223]	rs7329174 (G) [223]	0.28	
CSK	15q24.1	1.32	$1.04x10^{-9}$	EU	[224]	rs34933034 (A) [224]	0.27	
ITGAM	16p11.2	1.62	$1.61x10^{-23}$	EU, AS, HA	[153, 154, 162, 167, 185, 215, 219, 225-227]	rs9888739 (T) [153]	0.19	ITGAM R77H compromises leukocyte adhesion and phagocytosis
PRKCB	16p11.2	0.81	$1.4x10^{-9}$	AS	[228]	rs16972959 (A) [228]	0.23	
CLEC16A	16p13.13	1.23	1.34x10-8	AS	[229]	rs12599402 (A) [229]	0.36	
IRF8	16q24.1	1.16	$2.08x10^{-10}$	EU	[164, 191]	rs11644034 (A) [191]	0.17	
TYK2	19p13.2	1.20	$3.88x10^{-8}$	EU	[164, 185]	rs280519 (A) [164]	0.47	
IKZF3	17q21	1.9	4.83×10^{-9}	EU, AA, HA	[191]	rs8079075 (G) [191]	0.17	
ZPBP2	17q12	1.92	3.48×10^{-10}	EU, AA	[191]	rs1453560 (C) [191]	0.06	
CD40	20q12-q13.2	0.63	$2.0x10^{-8}$	EU	[230]	rs4810485 (T) [230]	0.24	
UBE2L3/HIC2	22q11.21	0.78	$1.48x10^{-16}$	EU, AS	[153, 155, 168, 169]	rs463426 (G) [168]	0.41	
TLR7	Xp22.3	0.60	$6.5x10^{-10}$	AS	[231-233]	rs3853839 (G) [231]	0.19	
IRAK1/MECP2	Xq28	1.39	$6.65x10^{-11}$	EU, AS, HA	[154, 162, 167, 234, 235]	rs1734787 (C) [235]	0.17	

[a] $P < 5x10^{8}$. [b] Genetic variants at 36 loci covering more than 50 genes have an established role in the etiology of lupus through association studies, replication in populations of diverse ancestry, and family studies. Some of the associated loci, such as 1q21 and Xq37, contain multiple genes. Until follow-up studies demonstrate or exclude the possibility of independent effects, all of the genes in the associated regions will remain candidate lupus genes. For HLA and complement, there are many different independent effects, which for simplicity, are summarized in a single

row. Variants in CFHR3/CFHR1, NMNAT2, ICA1, IKBKB, and SCUBE1[4] have *P* values approaching genome-wide significance ($5x10^{-8} < P < 10^{-7}$) and are the focus of current replication studies. Population ancestry: EU, European; AA, African; AS, Asian; HA, Hispanic American. FCGR, Low-affinity Ig- FcR; NCF2, neutrophil cytosolic factor 2; CRP, C-reactive protein; TNFSF, TNF superfamily[5]

To date, consensus surrounding the etiology and pathogenesis of lupus remains elusive. There is great potential in utilizing genetic studies to highlight signaling pathways that likely contribute to lupus pathogenesis. Genetic association data inform hypothesis-driven questions for functional biologists to elucidate underlying disease mechanisms.

In this review we present the many genetic regions now identified as associated with lupus. We then discuss how the putatively responsible genes are often co-expressed and cooperate in shared signaling pathways. Particular attention is focused on the convergence of lupus-associated variants in the B cell antigen-receptor (BCR) signaling pathway, as well costimulatory signaling pathways such as CD40, which synergizes with the BCR. The enrichment of genes in these pathways provides insight into SLE pathogenesis and offers potential targets for future therapeutic strategies.

Genetic associations with SLE

Over the past two decades, lupus genetic studies have transitioned from identifying linkage effects using variable number tandem repeats in less common multiplex families, to the statistical association of alleles defined by single nucleotide polymorphisms (SNPs). The standard study design for association studies compares allele frequency of SNPs (or other variant markers) in cases and controls. These studies can be performed as candidate hypothesis-driven studies on one SNP or a handful of

SNPs in one gene. Alternatively, in hypothesis-generating genome-wide association studies, the assessment of millions of human SNPs facilitates the discovery of statistically associated genetic variants that serve as markers for the variants that are causal. Presently, deep sequencing studies are underway which allow the exploration of the entire variant genotypic complement of an individual. The incredible technical revolution of the past seven years, analogous to switching from horseback to a modern sports car seemingly overnight, fundamentally changes the conceptual character of the genetic approach. Contemplating a complete identification of the major genetic associations and their interactions is daunting; however, given the current trajectory of innovation these types of studies might well be possible within the foreseeable future.

For lupus, fewer than 5% of the over 150 peer-reviewed and published associations are confirmed in the GWA era of genetic studies [236]. As the genotyping technology has become more sophisticated, the weak performance of the previous candidate gene and linkage eras became obvious for many disease related phenotypes [237, 238]; however, those genes that have been replicated such as IRF5 and STAT4 have modest effects and should not be discounted. In the spirit of a reemphasis on the scientific method and the development of more rigorous statistical criteria before an association would be accepted as probably present, Pe'er *et al.* developed a genome-wide criteria for significance in order to diminish the consequences of over-testing the data. They reasoned that the genome contains about 1 million independent elements, given the linkage disequilibrium present [239]. Therefore, the threshold for having a one in twenty chance of a false positive result (alpha error of 0.05) should be a probability of about $5x10^{-8}$ [239]. The problem is that this level of rigor requires large collections of

cases and controls to achieve adequate power with moderate expectations for effect size. Even then, the case-control design only has adequate power with a limited set of mechanisms of inheritance [240]. In fact, statistical power is very poor to detect recessive and epistatic effects, despite their presumed role in lupus etiology. The sample sizes required for most alleles operating through these mechanisms of inheritance are far beyond our current capacity. Notwithstanding the challenges, the application of these rules has allowed for the identification of more than 50 statistically convincing and replicated associations for the development of lupus (Table 4).

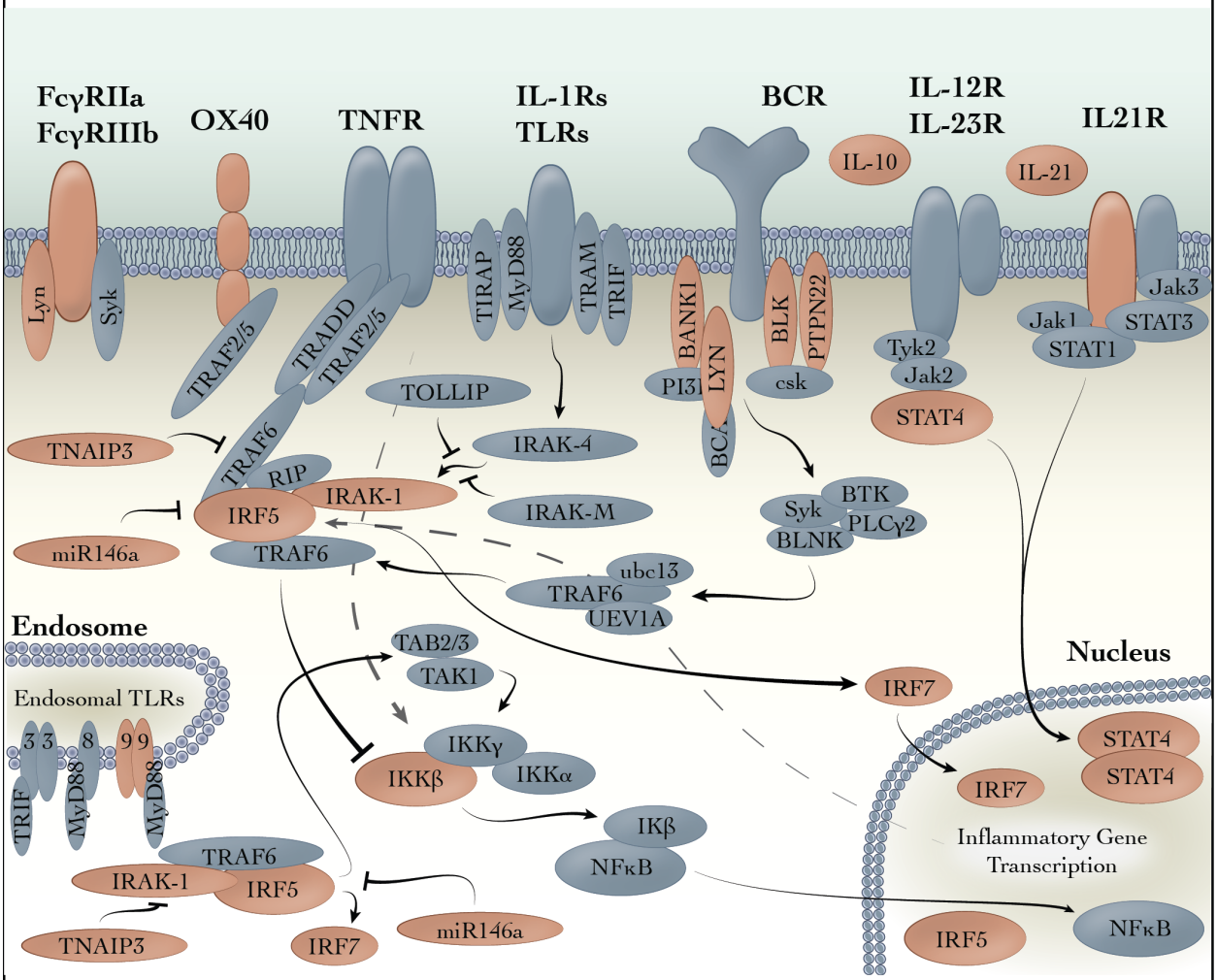
Once an association is established, fine mapping studies are initiated with the aim of identifying genetic variants that are biologically responsible for the genetic association. Fine mapping studies are usually done by designing genotyping arrays that use the same chemistry as the GWAS arrays to genotype as many SNPs as possible around the detected lupus-associated variant. Many such array-based fine mapping studies limit the location of the causal variant to small physical regions of DNA. The Large Lupus Association Studies (LLAS-1 and LLAS-2), encompassing 20,000 - 30,000 fine mapping markers and up to 18,000 subjects, have been effective for fine mapping identified loci (described in [143, 189, 241]). Subsequent studies currently underway with the 125,000 useful fine mapping markers on the ImmunoChip are expanding our ability to rapidly fine map associated regions, as has been recently accomplished for Celiac's disease [242]. Another, more recent, approach is to obtain deep DNA sequence of physical regions around the associated variants in a sample of subjects in order to provide a characterization of the variation differences between the risk and non-risk haplotypes.

Fine mapping studies of the strongest associations have revealed that identifying the causal variant is a non-trivial process for many genes. In fact, amino acid changing SNPs associated with lupus have only been identified and subsequently became candidate causal variants in 8 out of the 50 associated loci (Table 4). The linkage disequilibrium within associated regions results in many variants demonstrating statistical associations at a single locus. The evolutionarily diverse and four-fold longer population history of African ancestry makes possible important insights into variant identification that would not otherwise be possible from physical mapping experiments in Europeans; however, even when the genetic effect is localized to a few hundred bases, the functionality of the variant is not always directly apparent. While coding variants such as non-sense and mis-sense SNPs are easily annotated, variants that affect transcription factor binding sites, microRNA regulatory sites, RNA stability, or epigenetic regulation of a locus are more difficult to annotate given the current level of understanding.

Oftentimes, finding the functional effect of an association through experimentation with genotyped samples can inform the analytical determination of the causal variant. For example, rs9888739 results in an arginine to histidine amino acid change at position 77 in **ITGAM** that putatively compromises leukocyte adhesion, which at present is accepted as the causal variant responsible for the genetic association in the genomic region around **ITGAM** [243]. In other cases, such as **IRF5** and **TNFAIP3**, the association is due to the biological impact of multiple causal variants that affect the mRNA expression and stability [244, 245].

Organizing the associations into testable pathways provides a framework from which to ask hypothesis-driven, biologically relevant questions. In lupus, several groups have suggested collections of associated genes that work together in a common pathway and activated by common stimuli. For example, of the ~50 genes now known from published and preliminary data to be convincingly associated with SLE, it is striking that 30 are directly or indirectly involved in the NFκB signal transduction pathway related to IRF5 and IRF7 (Figure 2).

Figure 2: Lupus risk genes associated with the IRF-5 locus



IRF5 mediates interferon activation and inflammatory gene transcription via activation inflammatory gene transcription. IRF-5 is activated by a number of receptors, including endosomal TLRs and is downstream of a number of activating surface receptors including the TNFR, IL-12R and IL-21R. Following IRF-5 activation, IRF-5 translocates to the nucleus, where, as a transcription factor, it binds DNA to regulate gene expression.

Pathways in Lupus: Highlighting Genes in the B Cell Antigen Receptor-signaling Pathway

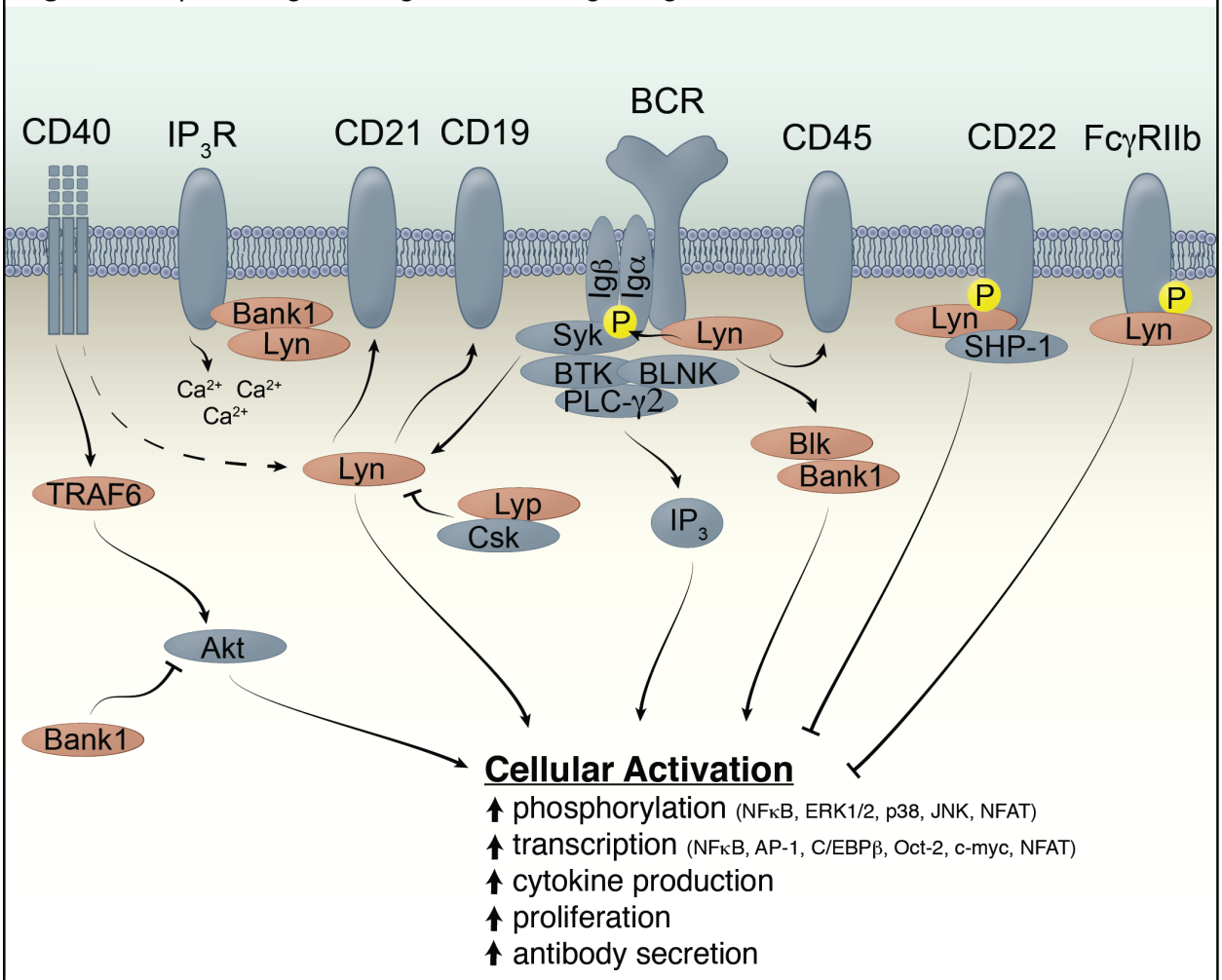
Reducing the complexity of disease-associated genetic variants by identifying cell types and pathways is a well-established approach to dissect lupus pathogenesis. Categorizing statistically associated genetic variants may reveal and highlight potentially important disease mechanisms. For example, Hu *et al.* uses a novel statistical method to identify cell specific expression of lupus variants and finds a convincing enrichment of gene expression in transitional B cell genes [246]. Lupus susceptibility genes tend to cluster in the Toll-like receptor (TLR)/type 1 Interferon (IFN) pathway, immune-complex processing, immune signal transduction, and B cell receptor (BCR) signaling pathway [247].

Intuitively, one might assume that multiple variants within the same category would synergistically increase risk of disease. Geneticists use complex statistical approaches to measure epistasis between variants in the same pathway. Despite studies including up to 15,000 subjects, the association results to date generally support additive relationships between lupus susceptibility genes instead of synergistic, multiplicative, or epistatic models [248]. In the context of current findings consistent with additive models of association, our current working hypothesis is that small changes to specific pathways permit the events that result in broken tolerance and overt autoimmunity.

To underscore significant lupus-associated genetic variants in B cell signaling pathways, we highlight the presence of proteins with established lupus associations in

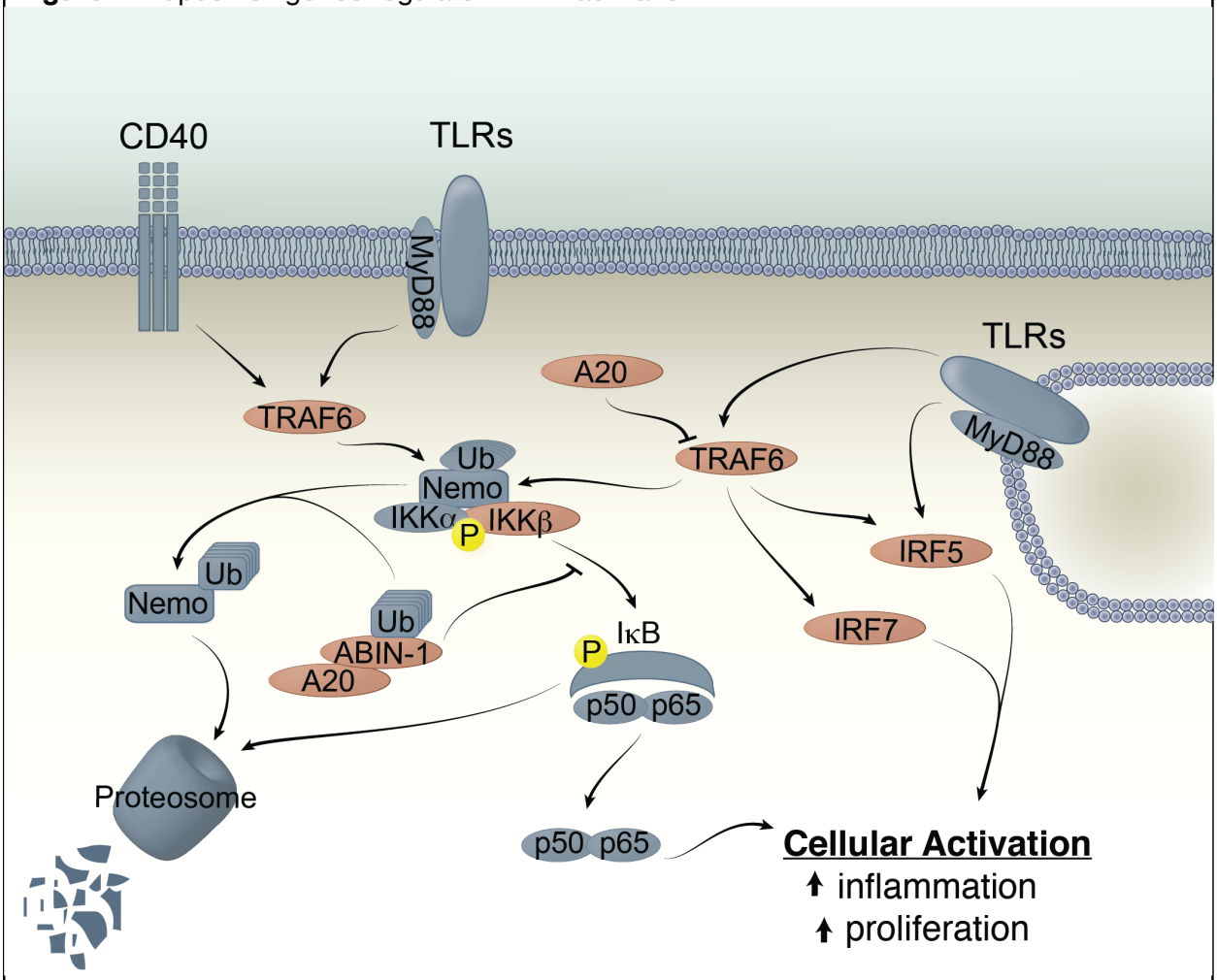
the BCR signaling pathway (Figure 3) and interconnected costimulatory pathways (Figure 4) with red shading in the figures and bold font in the text.

Figure 3: Lupus risk genes regulate B cell signaling.



Engagement of cell surface receptors, inhibitory or excitatory, induces signaling events involving numerous lupus-associated genes. LYN can bind to excitatory receptors, such the BCR, and inhibitory receptors, such as CD22 and FcγRIIb. In both cases, it acts as a kinase to initiate downstream activity, with the end result of activation or inhibition depending on the context of the receptor. BLK is also a PTK but does not interact with inhibitory receptors, resulting in the propagation of excitatory signals only. LYP is a PTP and thus, dephosphorylates second messengers in downstream activation pathways, resulting in reduced activation. BANK1 is a scaffold protein that brings PTKs into proximity with targets such as IP3R. BANK1 can also suppress cellular activation, for example, by inhibiting CD40-dependent Akt phosphorylation. Proteins shaded in brown are coded by lupus risk-associated genetic loci. NFAT, Nuclear factor of activated T cells.

Figure 4: Lupus risk genes regulate NF-κB activation..



Ubiquitination (Ub) is a key regulatory step for numerous cell-signaling events. A20 and TNIP1 form a complex that cleaves activating K63-linked ubiquitin chains (depicted in the figure by ovals) and add K48-linked ubiquitin chains (depicted in the figure by squares). Whereas the exact mechanism is not known, A20 and TNIP1 act in a MyD88-dependent manner to regulate inflammation, in part, by inhibiting TRAF6, which is an E3 ubiquitin ligase that is known to act downstream of CD40 and MyD88-linked extracellular (TLR1/2, TLR2/6, TLR4, and TLR5) and intracellular (TLR3, TLR7/8, and TLR9) TLRs. TRAF6 activation leads to the phosphorylation of IRF5 and IRF7, which among other functions, activate NF-κB signaling. Of note, each of these genes participates in many interactions beyond those depicted in this figure. Proteins shaded in brown are coded by lupus risk-associated genetic loci.

B cells in lupus

Patients with lupus progressively accumulate a number of autoantibodies directed against self-antigens, in particular anti-nuclear proteins[249]. While it is not clearly understood what leads to this break in self-tolerance, autoantibody production precedes overt disease manifestations and clinical diagnosis [46]. Autoantibodies are

thought to mediate many of the clinical sequelae of lupus, both indirect and direct. Renal disease in lupus is primarily caused by immune complex deposition in the glomerulus, leading to fibrosis and loss of glomerular filtration capacity [250]. Autoimmune hemolytic anemia is thought to be a result of clonal autoantibodies to cell surface antigens. Some lupus patients develop anti-phospholipid syndrome in which antibodies against the phospholipid-beta-2 glycoprotein I complex are associated with thrombotic events, placental infarcts leading to miscarriage, and liver dysfunction [251]. Given the role of autoantibodies in SLE pathogenesis, B cells have long been thought to play a crucial role in mediating disease.

Signaling through the B cell antigen receptor (BCR)

The BCR-mediated signaling cascade is a highly complex series of inhibitory and activating interactions that occur in spatial and temporal patterns (reviewed in [252, 253]). BCR-mediated signaling begins when antigen binds and cross-links the BCR. Src-family protein tyrosine kinases (PTKs), such as **Lyn**, phosphorylate immunoreceptor tyrosine-based activation motifs (ITAMs) on Igα and Igβ, which then serve as docking sites for cytosolic Src homology 2 (SH2)-domain-containing PTK spleen tyrosine kinase (Syk), resulting in phosphorylation of Syk (Figure 3). Upon activation, Syk is central to multiple downstream signaling events (PLCgamma2-SLP65/BLNK complex), resulting ultimately in cytoskeletal rearrangement, gene-expression regulation, and cell-fate decisions. As PTKs are central to positive BCR signaling, protein tyrosine phosphatases (PTPs), such as **Lyp,** encoded by **PTPN22,** play an opposing role and serve as a balance to maintain homeostasis. PTPs are

recruited to the BCR signaling complex through immunoreceptor tyrosine-based inhibitory motifs (ITIMs) on inhibitory co-receptors such as CD5 and CD22 (Figure 3). ITIMs are also targeted by Src-family PTKs, however, ITIM phosphorylation results in recruitment of SH2-domain containing PTPs, which counteract the kinase activity of PTKs. Another mechanism by which BCR signaling is negatively regulated is by activation of the low-affinity Fcγ receptors (FcγRs), specifically **FcγRIIb**. This receptor is unique among FcγRs in that it has an ITIM, and phosphorylation of this motif results in attenuation of BCR signaling (Figure 3).

Stimulation of B cells on coreceptors such as **CD40** or through **TLR** activation synergize with BCR signaling and lead to B cell proliferation (Figures 3 and 4), cytokine production, antibody secretion, and isotype switching, each of which play potential roles in SLE pathogenesis. Different combinations of costimulatory signals allow B cells to produce cytokines and antibody isotypes for a specific inflammatory response. Genetic variants in these costimulatory pathways subtly affect how B cells generate an immune response. These changes could prejudice B cell responsiveness toward inappropriate breaks in tolerance leading, in some cases, to lupus.

One distinct feature of the BCR signaling pathway is that in addition to second messenger mediated signal transduction, the BCR-antigen complexes are physically internalized and trafficked through the endocytic pathway to the MHC class II-containing compartment (MIIC). Here antigen is digested and loaded onto MHCII molecules prior to surface expression [254].

BCR signaling in B cell selection

The mechanisms that allow auto-reactive B cells to escape negative selection and become antibody-secreting plasma cells in the periphery are poorly understood. It is clear that tolerance is broken in these auto-reactive B cells and that immune dysregulation is central to the etiology of lupus. Currently, we do not have clear mechanistic insight into how genetic and environmental factors sometimes result in the loss of immunological tolerance of self. B cell selection occurs in two general steps. Early selection prior to development into mature B cells, and later a second selection step occurs as mature B cells circulate in the periphery. Early B cell selection consists of two checkpoints. First, immature B cells in the bone marrow (BM) can undergo three distinct fate decisions based on binding affinity to self-antigen; high-affinity binding results in deletion, moderate-affinity binding leads to receptor editing via heavy and light chain gene rearrangement while low-affinity binding leads to anergy. The second checkpoint is less well understood, but is believed to occur as newly emerged transitional B cells circulate in the periphery prior to maturation. This is supported by the decreased numbers of autoreactive clones in the mature B cell pool when compared to the number of autoreactive clones in the pool of new-emigrant B cells, suggesting that some deletion of auto-reactive B cells has occurred in transit. Many suspect that this may result from a new-emigrant B cell's encounter with autoantigen leading to subsequent negative selection [255].

Specific genes with a potential role in B cell signaling and survival

LYN

LYN encodes a Src-family PTK that acts early in the B cell signaling pathway (Figure 3). Polymorphisms within **LYN** are associated with lupus in European-American women and the risk allele is the major allele [142, 256]. The specific genetic variant responsible for the lupus association with **LYN** is unknown though **Lyn** expression is decreased in B cells isolated from lupus patients. The decrease in expression also correlates with exclusion of **Lyn** from the lipid raft fraction in B cell digests [257, 258]. **Lyn** is one of the central mediators of signaling in B cells as it phosphorylates both ITAMs and ITIMs (Figure 3).

Early results with **LYN**$^{-/-}$ mice revealed a lupus-like phenotype that suggested a role for **Lyn** in lupus pathogenesis. **LYN**$^{-/-}$ mice have decreased B cells in the periphery that display decreased proliferation and diminished phosphorylation of CD22 and **FcγRIIb** in response to anti-IgM stimulation [259]. Autoantibody levels are increased in **LYN**$^{-/-}$ animals, as is total IgM in the serum, which corresponds with an increase in the IgM+ B cell fraction [260]. While **Lyn** can phosphorylate both ITAMs and ITIMs of cell surface receptor complexes, its role in activating positive signaling pathways downstream of the BCR is redundant with **Blk** and Fyn [261]. Importantly, **Lyn** has a unique role in mediating negative signaling by phosphorylation of ITIMs in CD22 and **FcγRIIb** [262, 263]. In an effort to understand the mechanism of the B cell depletion in **LYN**$^{-/-}$ mice, Gross *et al.* generated mixed bone-marrow chimeras to define the competitive advantage or disadvantage of **Lyn** deficiency in B cells. The loss of **LYN**$^{-/-}$ B cells is a B cell intrinsic defect, as **LYN**$^{-/-}$ follicular B cells were reduced compared to wild type in mixed bone marrow chimeras. Mechanistically, this B cell contraction correlates with increased expression of the pro-apoptotic molecule Bim [264].

As **Lyn** can act as both a positive and negative regulator, **LYN**$^{-/-}$ mice are used to investigate the balance between positive and negative signaling in peripheral B cells. Newly arrived splenic transitional (T1) B cells are progressively less sensitive to BCR-induced signaling as they mature to follicular B cells. This is demonstrated by increased Ca^{2+} flux and second messenger phosphorylation in T1 B cells compared to follicular B cells following IgM stimulation. The increased signaling is most likely due to the upregulation of the CD22-SHP-1-**Lyn** regulatory axis, which is an important source of negative signaling (CD22 is increased in murine follicular B cells compared to T1 B cells). T1 B cells from **Lyn**-deficient mice show no change in BCR induced signaling compared to wild type, while **Lyn**-deficient follicular B cells are much more active than wild type cells [265]. These results further suggest a role of **Lyn** in maintaining negative signaling in B cells, particularly follicular B cells. While these results highlight a role for **Lyn** in the suppression of follicular B cell activation, loss of **LYN** does not affect tolerance during B cell maturation as **LYN**$^{-/-}$Ig^{HEL} mice have no defect in negative selection [266]. These data identify **Lyn** as a key mediator of the **FcγRIIb** and CD22 negative signaling pathways, which appear to be more important in mature B cells compared to early transitional cells, where stimulatory signals predominate, promoting negative selection. The increase in positive signaling observed in **LYN**$^{-/-}$ B cells may be enough to increase negative selection in early transitional B cells, leading to the decreased pool of transitional B cells in **LYN**$^{-/-}$ mice. The explanation for the emergence of autoreactive B cells and autoantibodies is not as clear given the decrease in B cells. Increased BCR signaling as a result of the decreased negative signaling in mature B cells due to **Lyn** deficiency is one explanation, but it may also be a product of the effect

of **Lyn** deficiency in myeloid cells that increases the inflammatory tendency of the cellular milieu [267]. **LYN**$^{-/-}$MyD88$^{-/-}$ mice do not produce autoantibodies [268] due to the suppression of plasma cell differentiation. The development of autoimmunity in **LYN**$^{-/-}$ animals most likely arises as a result of the combination of the increased sensitivity of mature B cells to positive signaling pathways and the concurrent increase in inflammation mediated by **LYN**$^{-/-}$ myeloid cells and TLR-dependent stimulation.

- **Lyn is a Src-family PTK that can phosphorylate both ITAMs and ITIMs (Figure 3)**
- **Lyn-mediated phosphorylation of ITIMs is crucial to dampen activation of follicular B cells in mice**
- **Decreased levels of Lyn are associated with lupus**
- **The major allele in the population is the risk variant for LYN**

BLK

Blk is a Src-family PTK expressed predominately in B cells [269]. The genetic association between **BLK** and lupus in individuals of European descent was first identified at a SNP ~7kb upstream of **BLK** and confirmed in subsequent studies [142, 270]. BLK expression is genotypic with people who carry two copies of the risk allele expressing more **Blk** than people carrying one copy or two copies of the protective allele [270]. The association of **BLK** with lupus has been replicated in Japanese [271] and Chinese [270] populations.

BLK interacts with **BANK1** genetically (with epistasis between the risk variants despite no increased lupus risk) and physically (by co-localizing) (Figure 3) [272]. Co-

immunoprecipitation studies demonstrated a basal interaction between **Blk** and **BANK1** that was enhanced by IgM stimulation. Additionally, the expression of **BANK1** appears to alter **Blk** localization, with **Blk** localizing to cytoplasmic compartments instead of the plasma membrane in the presence of **BANK1** [272].

BLK$^{-/-}$ deficient mice have no observable immune phenotype [273]; however, because **Blk** is functionally redundant with **Lyn** and Fyn, this is not surprising [261]. While **Blk** does not apparently have a unique role in mice, decreases in **Blk** expression may still have a significant impact on human disease, as is seen with another PTK family gene product, Bruton's tyrosine kinase (Btk). Mutations in Btk in humans results in X-linked agammaglobulinemia, which can result in a complete lack of B cells. The phenotype in BTK$^{-/-}$ mice is not as severe, suggesting significant differences between PTK signaling pathways in humans and mice [274].

- **Blk is a PTK that can bind both ITAMs and BANK1 (Figure 3)**
- **A polymorphism in BLK associated with lupus is also associated with decreased expression of Blk**

BANK1

B-cell scaffold protein with ankyrin repeats 1 (**BANK1**) is a B cell scaffold protein that binds both the Src family PTK **Lyn** and the inositol 1,4,5-triphosphate receptor (IP_3R) in distinct regions. Upon BCR stimulation, **BANK1** is phosphorylated, favoring **Lyn** binding (Figure 3). Dual binding of **Lyn** and IP_3R permits **Lyn**-mediated phosphorylation of IP_3R, leading to Ca^{2+} mobilization from intracellular ER stores [275]. **BANK**$^{-/-}$ mice have been developed and found to have slightly increased numbers of mature B cells, increased

serum IgG2a and increased formation of germinal centers. Following CD40 stimulation, B cells from **BANK1**-deficient mice demonstrate enhanced proliferation and survival. **BANK1**$^{-/-}$,CD40$^{-/-}$ mice rescue the phenotype. The increased survival and proliferation of B cells depends on CD40 activation of TRAF6 which is essential for Akt ubiutination, membrane recruitment , and phosphorylation. Further mechanistic characterization of the BANK1-/-CD40-/- mice suggest that **BANK1** acts as a negative regulator of CD40-induced survival signals via the suppression of Akt [276]. The connection between the role of BANK1 in Lyn-mediated signaling and the role of BANK1 in CD40 mediated signaling within B cells remains to be explored, as does the possibility that BANK1 has actions that influence the BCR and CD40 pathways distinctly at different times over the life of a B cell. Future studies will also need to address the role of BANK1 in B cell signaling pathways apart from CD40 and Lyn. Two major isoforms are identified for **BANK1**: one full-length isoform and a truncated isoform lacking all of exon 2, named Δ2 [277].

Kozyrev *et al.* identified three potential functional variants associated with lupus in Europeans, a finding which we subsequently replicated [277, 278]. rs17266594 is in a putative branch point site in intron 1, but was later shown not to be essential for splicing [279]. rs10516487 codes for an R61H variant in exon 2, and rs3733197 which codes a A383T variant in exon 8 [277]. The non-synonymous SNP rs10516487 is located in the second exon and is in high linkage disequilibrium with rs17266594 (r^2=0.90). The protective minor allele A results in a coding change from Arg to His at position 61, resulting in the creation of a recognition site for the splicing enhancing factor SRp40. Presence of this binding site appears to inhibit splicing of exon 2. In addition, the A

allele was associated with lower levels of **BANK1** expression at both the message and protein level. The exact relationship between the R61H variant and the Δ2 isoform are not firmly defined, but the Δ2 isoform lacks the IP_3R binding domain and when present either alone or with full-length **BANK1**, is diffusely distributed in the cytoplasm instead of clustered into punctae as is the full length isoform when present alone. Thus, heterologous dimer formation may be an important mechanism of **BANK1** regulation, determining cellular location and access to interacting proteins [279]. While the identification of these potential causal variants is promising, the exact relationship between SLE-associated variants in **BANK1** and disease pathogenesis remains unclear. Further studies focused on understanding the effect of these functional variants on the CD40 signaling pathway will help clarify the role of **BANK1** in SLE pathogenesis.

- **BANK1 is a scaffold protein that can bind PTKs and IP_3R (Figure 3)**
- **Multiple functional polymorphisms in BANK1 are associated with lupus, although the casual variants are not yet identified**
- **B cells from BANK1-deficient mice proliferate in a CD40-dependent manner**

PTPN22

PTPN22 encodes **Lyp**, a PTP found in lymphocytes [280]. A variant of **PTPN22** (1858C/T) is associated with multiple autoimmune diseases, including RA [281], SLE [282], T1D [283], Graves disease [284] and myasthenia gravis [285]. The association between **PTPN22** and SLE is well established in multiple ancestries [286] and seems to be entirely explained by the 1858C/T polymorphism. This genetic variant results in the amino acid substitution R620W, producing **Lyp** 620W.

Lyp interacts with Csk, a PTK family member, to regulate phosphorylation of Src family kinases downstream of cell-surface receptors including the TCR, BCR and **CD40** (Figure 3). Early studies demonstrated that the R620W substitution disrupts **Lyp** binding to Csk [283]. The **Lyp** 620W variant has been studied in lymphocytes isolated from patients with lupus, T1D, and myasthenia gravis. The functional consequences of the **Lyp** 620W variant in autoimmune disease were first reported in T cells isolated from T1D patients [283, 287]. In these cells, TCR engagement resulted in decreased phosphorylation of downstream targets, decreased Ca^{+2} flux, and decreased IL-2 production [287]. Given these findings, the **Lyp** R620W mutation was proposed to be a gain-of-function mutation, resulting in decreased signaling through the TCR in cells carrying the disease-associated phenotype [287]. Subsequent work in isolated T cells appears to have confirmed these findings [288, 289]. However, other groups have reported the opposite phenotype in both peripheral blood mononuclear cells (PBMCs) and T cell lines simultaneously transfected with both **Lyp** and Csk [222, 290]. While T cells isolated from individuals carrying the **Lyp** R620W mutation may represent a more physiological assessment than overexpression of **Lyp** *in vitro*, these differences remain to be reconciled.

Current efforts to resolve these differences have focused on the mouse homologue of **Lyp:** PEP. Zikherman *et al.* crossed $PEP^{-/-}$ mice with $CD45^{w/w}$ mice, which contain a point mutation in the CD45 juxtamembrane wedge domain (E613R), in which hyperactive CD45 signaling leads to B cell-driven autoimmunity [222]. The $PEP^{-/-}$ $CD45^{w/w}$ animals develop a lupus-like disease, similar to $CD45^{w/w}$ mice, but with increased anti-double stranded DNA IgG. B cells from $PEP^{-/-}CD45^{w/w}$ mice are similar to

those from CD45$^{w/w}$ mice. While attempts to understand the effect of PEP deficiency on the development of autoimmunity have not been impressive, to date, other mouse models have been developed to specifically test the hypothesis that the **Lyp** 620W mutation will lead to increased autoimmunity. To test this hypothesis, mice expressing a PEP variant similar to **Lyp** 620W, PEP 619W, were developed. Expression of the **Lyp** 620W/ PEP 619W variant results in lower **Lyp**/PEP protein expression due to increased calpain- and proteasome-mediated cleavage and degradation, causing an increased activation of lymphocytes in these mice [291].

The Bruckner lab has produced numerous studies over the last few years examining the phenotype of B cells isolated from human subjects carrying the 1858C/T **PTPN22** polymorphism. Individuals carrying at least one copy of the risk allele have reduced memory B cells in the periphery [289], along with increased transitional and anergic naïve B cells [292]. Strikingly, autoreactive B cell clones are more frequent within these pools of transitional and mature naïve B cells. These B cells are also more sensitive to CD40L and IgM crosslinking by measure of surface activation markers [293]. In contrast to what occurs in the transitional and anergic naïve B cell pools, the memory B cell population of individuals carrying the lupus-risk allele demonstrate a decrease in Ca^{2+} mobilization and proliferation following treatment with anti-IgM *in vitro*. In addition, stimulation of these memory B cells results in a reduction in phosphorylation of downstream targets, including Syk and PLCy. The genotypic attenuation in these signaling pathways can be restored following treatment with the **Lyp** inhibitor I-C11 [294]. These studies suggest that **PTPN22** plays a crucial part in early B cell tolerance checkpoints, with **Lyp** 620W permitting progression of autoreactive clones to the

periphery. These results also highlight the B cell-intrinsic effects of **Lyp;** however, **PTPN22** mutations in T cells may still be important in the modulation of B cell activation. Decreased signaling through the TCR could also affect Treg function, thereby allowing autoreactive B cells that escaped to the periphery as the result of **PTPN22** mutations during selection to be activated and drive SLE pathogenesis.

If **Lyp** 620W does represent a gain-of-function mutation, then one explanation for the phenotype observed in individuals with the 1858C/T polymorphism would be that a hyperfunctional **Lyp** restricts activation of key second messengers, as evidenced by decreased phosphorylation in cells with **Lyp** 620W, raising the threshold for cellular activation. A higher activation threshold in transitional and naïve B cells could allow persistence of B cell clones that would otherwise be deleted through high-affinity binding. This is reflected in the increase of autoreactive cells in the pool of transitional B cells and importantly, may be a key event in the breakdown of tolerance and progression towards autoimmunity as the increase in autoreactive clones mirrors what occurs in Type 1 Diabetes (T1D) [293].

In an additional layer of complexity, gene expression analysis identifies **IRF5** as one of the genes upregulated in B cells from 1858C/T individuals [293]. **IRF5** is well established as being an important SLE risk gene. Increased **IRF5** expression is associated with Type I IFN production [295], suggesting that the **PTPN22** risk allele may also promote cellular activation in the periphery. This represents a further mechanism by which alterations in which **PTPN22** polymorphisms promote autoimmunity.

- **PTPN22 encodes Lyp, a PTP important for negative regulation of B cell signaling (Figure 3)**

- **A PTPN22 risk allele, the Lyp 620W putative casual variant, is associated with numerous humoral autoimmune diseases**
- **Lyp 620W is associated with increased activation of transitional and follicular B cells and decreased activation of memory B cells**

TNFAIP3

Tumor necrosis factor alpha inducible protein 3 (**TNFAIP3**), encoding the ubiquitin-modifying protein **A20**, was first identified as a risk gene in RA [296, 297] and has since been associated with multiple autoimmune diseases, including lupus [298, 299], Celiac disease [300], Crohn's disease [301], T1D [302] and psoriasis [303]. Recently, a TT>A deletion polymorphism was proposed to be the causal variant responsible for disease association in lupus. This deletion is downstream of the **TNFAIP3** gene in an apparent regulatory region. Oligonucleotides with the deletion display decreased binding of the NFκB subunits p50, p65 and cREL and reduced overall **TNFAIP3** mRNA expression according to the number of risk chromosomes (0, 1 or 2) present in the sample [304]. These results suggest that decreased **A20** levels are a risk factor for lupus.

A20 is an ubiquitin-editing enzyme induced by CD40 activation and signaling by proinflammatory cytokines such as TNFα and IL-1β. **A20** modulates cell signaling by both removing K63-linked ubiquitin chains from signaling intermediates and by adding K48-linked ubiquitin chains [305]. In general, the addition of K63-linked ubiquitin chains is activating while K43-linked polyubiquitination targets proteins for proteosomal degradation. The main byproduct of ubiquitin editing events is to limit NFκB activation,

aiding in the maintenance of immune homeostasis (Figure 4) [306]. Recent evidence suggests that inhibition of IKKβ phosphorylation by **A20** may be independent of **A20** enzymatic activity, though still dependent upon ubiquitin binding [307]. **A20** deficient mice develop massive systemic inflammation resulting in cachexia and multi-organ failure leading to early death mediated by failure to restrict NFκB signaling [308]. Crossing **A20** $^{-/-}$ mice with MyD88$^{-/-}$ mice rescues the phenotype by restricting ubiquitination of **TRAF6** [309], suggesting that **A20** plays a prominent role in dampening homeostatic MyD88-mediated inflammatory signaling (for a recent review see [310]).

Mice with **TNFAIP3** conditionally deleted in B cells (CD19+) were recently developed by two separate groups [311, 312]. Both report significant changes in B cell populations, including increased plasma cells in older mice, immune complex deposition in the kidney and increased IL-6 production. Tavares *et al.* reports significant autoantibody production [312], but it is not clear if these are IgM or IgG autoantibodies. Since some IgM autoantibodies may be protective [313], the implications of this finding are not immediately obvious. The **TNFAIP3**-deficient mice studied by Chu *et al.* had high levels of IgM in the sera and class-switched anti-cardiolipin IgG in older mice [311].

Furthermore, both groups also report changes relevant to splenic B cells in the B cell conditional TNFAIP3$^{-/-}$ mice. Chu *et al.* finds that marginal zone B (CD1d+) cells are located in the follicle of the spleen, with little to no B cells in the marginal zone. In addition, they demonstrate that splenic B cells fail to differentiate into Blimp1+ plasmablasts following LPS stimulation *in vitro* [311]. Tavares *et al.* also identify alterations in splenic B cells, noting an increase in germinal center B cells [312].

To test the hypothesis that **A20** plays a role in Fas-mediated programed cell death in germinal center negative selection [314], programed cell death was induced by agonist anti-CD95 stimulation *in vitro*. Indeed, **A20** deficient cells are resistant to Fas-mediated programmed cell death. The resistance to Fas-induced apoptosis correlates with an increase in the anti-apoptotic molecule Bcl-x, a molecule that is normally suppressed by the NFκB pathway [311]. Notably, both groups observe a dose effect in response to **A20** deficiency in B cells, consistent with a pathogenic role for alleles producing less active **A20** leading to B cell hyperreactivity in human autoimmunity.

Taken together, **A20** is potentially critical to the homeostatic dampening of NFκB signals and that decreased protein expression is sufficient to tip B cells toward an autoimmune phenotype and loss of tolerance. While the exact mechanism remains to be identified, the results from conditional knockout mice suggest that decreased expression of **A20** may result in the production of B cells resistant to negative selection in the spleen. In addition, increased IL-6 produced by these animals may also promote the survival of autoreactive plasma cells and T cell mediated pathogenesis. Obviously much work remains to connect these murine phenotypes with the genetic associations reported in human SLE. However, given the findings that SLE-associated genotypes correlate with lower **A20** expression and that decreased **A20** in murine B cells results in resistance to negative selection, *in vivo* studies aimed at tying together these observations will be an exciting next step in using B cell signaling pathways to understand statistical genetic associations.

- **TNFAIP3 encodes A20, a ubiquitin editing enzyme that can both add and remove ubiquitin chains from target molecules (Figure 4)**

- **A20 can regulate inflammation via attenuation of MyD88-dependent NFκB signaling in mice**
- **A lupus-associated proposed causal variant in TNFAIP3 is statistically associated with lower expression of A20**

TNIP1

TNIP1 encodes an ubiquitin-binding adapter protein **A20**-binding inhibitor of NFkB, **ABIN-1**, that binds and cooperates with **TNFAIP3** in suppressing NFκB (Figure 4). Also known as NAF1, polymorphisms in **TNIP1** are associated with lupus in both European derived and Asian populations [143, 315]. **TNIP1** can be differentially spliced into at least 11 different isoforms [316] each of which suppress NFκB activation with different levels of inhibition. Whether the lupus-associated risk variant of **TNIP1** alter splicing or NFκB activation is not yet known.

Deubiquitination and degradation of ubiquitinated signaling proteins are key mechanisms by which **ABIN-1** cooperates with **A20** to limit T lymphocyte receptor (TCR) signaling as part of a negative feedback loop. Coexpression of **ABIN-1** together with **A20** in HEK293 cells contributes to the negative regulatory function of **A20** in TNF-α–mediated NFκB activation [317]. Mechanistically, **ABIN-1** interacts with **A20** independent of the ubiquitin-binding domain to mediate the NFκB inhibitory function of **A20** [318].

ABIN-1 also binds the MAP kinase ERK2 and causes retention of ERK2 in the cytoplasm, leading to reduced signaling [319]. siRNA mediated knockdown of **ABIN-1** results in increased IFN-β production following either viral infection or polyI:C

administration *in vivo*, which is dependent upon the ubiquitin-binding domain of **ABIN-1**, suggesting that **ABIN-1** suppresses viral-induced IFNβ production [320]. Interaction between **ABIN-1** and the toll-like receptor (TLR) pathways is further supported by the finding that **ABIN-1** recognizes K63-linked polyubiquitin chains on NEMO. NEMO is the IKKγ subunit of the NFκB signaling complex. NEMO activation is necessary for the nuclear translocation of NFκB. While K48-linked polyubiquitination of NEMO targets it for subsequent proteasomal degradation K63-linked polyubiquitination of NEMO regultates its activity, but not half-life [321].

TNIP1-deficient mice develop progressive, lupus-like inflammatory disease. The development of the lupus-like disease is dependent on leukocytes, as bone marrow transplantation of wild type mice with **TNIP1**$^{-/-}$ bone marrow is sufficient to initiate autoimmunity. **ABIN-1** is recruited to the MyD88 signaling complex and controls TLR-mediated activation of the major isoform of "CCAAT enhancer-binding protein b" (C/EBPb) which is known as LAP (liver-enriched activator protein) which has been shown to mediate TLR-dependent inflammatory cytokine production [322]. In fact, mice carrying a mutant **TNIP1** that cannot bind ubiquitin also display an autoimmune phenotype. Ubiquitin binding is rescued by crossing the mice onto a MyD88$^{-/-}$ background [323]. Stimulation with TLR ligands, BCR crosslinking, or a **CD40** agonist all caused enhanced phosphorylation and activation of the NFκB pathway in **ABIN-1** mutant mice compared to wildytpe mice [323]. These murine studies lead us to conclude that **ABIN-1** cooperates with **A20** in BCR signaling through interaction with the signaling pathway components downstream of both **CD40** and the BCR.

We recently reported a potential genetic association between SLE and another protein involved in ubiquitination, **TRAF6** [324]. **TRAF6** is an E3 ubiquitin ligase that has been identified as being important in numerous immune functions by activating NFκB following stimulation of BAFFR, TNF receptors, TLRs and **CD40** [325, 326]. While the results of these first association studies will need to be replicated in many more people to confirm this association, they are consistent with the reported association of **TRAF6** with rheumatoid arthritis, also a humoral autoimmune disease. These results emphasize the importance of ubiquitin-mediated modulation of cellular activation pathways. The associations between these genes and lupus suggest a fruitful area of research in working to elucidate the mechanisms by which disruptions in ubiquitination can facilitate the onset of autoimmunity (See Figure 4).

- **TNIP1 encodes a ubiquitin-binding adapter protein, ABIN-1, that binds both ubiquitin and A20 (Figure 4)**
- **Several genes involved in the ubiquitin regulatory pathway (e.g., TNIP1 and TNFAIP3) are associated with lupus highlighting the importance of this pathway in the development of autoimmunity**
- **ABIN-1 has anti-inflammatory activities that are both dependent and independent of ABIN-1 ubiquitin binding in mice**
- **ABIN-1 attenuates MyD88/NFκB inflammation in mice in part by targeting NEMO for proteasomal degradation; furthermore, ABIN-1 acts independently of MyD88 downstream of BCR and CD40 signaling.**

Concluding Remarks

B cells are gaining newfound respect, now being appreciated as more than immunoglobulin factories. Indeed, the genetic evidence supplements other data supporting their critical role. Early case reports for small clinical studies using B cell depletion therapies showed promise that these strategies would prove to be highly effective in treating lupus [91, 327, 328]. Unfortunately, larger random control studies did not replicate these findings [95, 329]. Similarly, the first drug approved for lupus in 50 years, belimumab, targets BAFF, a cytokine essential for B-cell survival and proliferation, and has shown modest benefit in clinical trials [330]. These results underscore the fact that there are other factors that contribute to inflammation and SLE pathogenesis. These targeted approaches will most likely be more effective in a subset of SLE patients. B cell-related genetic risk variants may one day be used in more sophisticated and successful clinical approaches, including identifying which patients will be most likely to respond to specific therapies.

We are early in the effort to understand how sequence changes to specific genes can affect lupus risk. As each of the lupus-associations is progressively evaluated, many complex, interconnected and sometimes contradictory stories are likely to emerge. The controversies and apparent contradictions in studies of **PTPN22** and **LYN** may foreshadow potential contradictions as we learn more about less studied associations such as **BANK1**. In fact, a decade ago **PTPN22** was thought to have a mechanism that operated through T cells. Now the evidence is just as strong that lupus risk related to **PTPN22** is operating through B cells. The history surrounding studies that ascribe molecular mechanisms to explain genetic effects gives us a glimpse at the

challenges of forming a reasonable conceptual framework to understand how risk of disease is generated.

Epstein-Barr virus (EBV) may play a significant role in the etiology of lupus. Briefly, the very earliest autoantibodies to appear in pre-clinical lupus patients are cross-reactive with EBV proteins, including EBNA-1, an EBV-encoded molecular mimic of SLE-associated autoantigens such as Sm and Ro [46, 331]. In addition, EBV seroconversion [332] and viral load are increased in lupus cases compared to controls [333]. Several of the BCR-associated lupus genes discussed in this review play a role in EBV replication and immune evasion. For example, **LYN** has been shown to be the preferred signaling protein for the major EBV latency protein latent membrane protein 2A (LMP2A), which acts as a BCR functional mimic, permitting B cell survival and preventing apoptosis of EBV infected B cells at least in part by co-opting the B cell signaling apparatus, including lipid rafts [334]. **A20** has also been shown to be important in EBV infection of B cells, as **A20**-mediated ubiquitination of **IRF7** is stimulated by the EBV protein latent membrane protein 1 (LMP1) [335], which functions as a **CD40** functional mimic to stimulate B cells. **A20** competes with the LMP1 complex in activation of the NFκB pathway, altering the ability of LMP1 to drive NFκB activation.

The biological impact of lupus-associated genetic variants in genes known to be important for EBV infection such as **LYN** and **A20** will be of special interest moving forward, as an individual's immune response to EBV infection (and the immune dysregulation caused by EBV-encoded mimics of key B cell activators) may provide important insight into EBV-associated lupus pathogenesis.

This review focused on the signaling in B cells and the corresponding lupus risk genes (Figures 3 and 4). Other aspects of B cell signaling, including BCR internalization and delivery of antigen to the MIIC are also known to have major roles in regulating B cell responses [336]. Genes in these pathways remain likely candidates for lupus. Also, other lupus genes, such as **STAT4** and **IRF5**, are important transcription factors that mediate the response of many different cells, including B cells, to inflammatory stimuli. Yet other genetic associations are in genes for which little is known. Further investigation might well reveal newly recognized proteins with critical B cell functions. As more sophisticated techniques become available, we may find that risk polymorphisms in separate pathways act in concert to facilitate cellular changes conducive to tolerance breakdown and autoimmunity.

Indeed, all the known genes together only account for a minority of the genetic heritability, as has been discussed for complex genetic diseases in general [337]. Together, small changes in proteins caused by common genetic variants alter B cell behavior toward a breakdown in tolerance and autoimmunity. Even after each of the individual associations are mechanistically understood, the practice of identifying the etiological mediators of lupus in patients presenting in clinic will require a sophisticated understanding of the interactions of the proteins in molecular pathways, such as the BCR signaling cascade and the regulation of ubiquitin modification. By integrating genetic and biological data through multi-disciplinary studies, we will prime ourselves for success in the coming age of personalized genomic medicine.

Chapter 3: Approach to Data Processing and Statistical Analysis in Genetic Association Testing

Introduction

The problems of Big Data analysis are widespread across many industries, from science to advertising to market research. The field of genetics is currently experiencing a period of great discovery, powered in large part by the increased capacity to query the genome. With the advance in technology available to the genetic researcher and the subsequent massive increase in data, there has been a rapid development of new tools specifically tailored to address the analysis needs of the genetic research community. There exists now a plethora of very specific tools, each addressing specific data types or analysis strategies. Whether beginning with processing raw sequencing or SNP-chip data to performing advanced statistical analysis of multiple data sets using different statistical approaches, there are most likely multiple options available for each step. Many of these tools are developed by bioinformaticians and computer scientists with specific interests in genetic analysis and software development. The tools they produce are most often open source to the research community and are continually updated and supported by their original creators.

While these programs can be quite powerful, they are often quite narrow in focus and complete analysis of any project requires utilizing multiple tools developed by a variety of labs. This creates a challenge to even an experienced bioinformatician as many of these programs were developed for different computer operating systems, have

sometimes little to no instructions on use and getting support from the developers can be problematic. For the genetic researcher without a background in computer science the learning curve can be prohibitively steep. The emergence of large software suites created by independent software development companies is an attempt to address these obstacles. They are designed to include many different types of analysis tools and are user friendly but come at a steep cost, with individual licenses running into the thousands of dollars a year.

Another obstacle presented by the advent of large genetic data sets is the amount of computer resources needed to perform a given analysis often exceeds the resources of an individual workstation. In addition, some analysis jobs can take many hours to days to complete, making it impractical to run on an individual user's machine. To overcome this limitation of processing power and time constraints, many institutions have developed large computational clusters consisting of hundreds of processing cores which can used to run analysis jobs. For example, the CCHMC computational cluster (bmi.cchmc.org/resources/clusters) has over 700 processing cores and runs under a Linux-based operating system. Users can access the cluster from both on and off-campus IP address. Jobs can be submitted and queued until the computational resources requested are available on the cluster. An individual user can also submit multiple jobs simultaneously, greatly increasing efficiency.

Given the specific needs of our lab to be able to analyze various types of genetic data reliable and consistently, with the need to consistently and rapidly repeat the analysis with new data sets as available, we developed a series of best practices to

ensure that we were consistent with the genetic analysis we performed on each data set and of equal importance, that even the newest member of the lab, with some initial introduction and previous moderate familiarity with computers would be able to produce reliable and repeatable data quickly. To facilitate reproducibility and ease of use, we also developed a series of Linux bash scripts, essentially short computer programs, which can be used to submit analysis jobs to the CCHMC computational cluster in a simple and repeatable manner. The data set types we chose to focus our initial effort on were large SNP genotype databases that are used in association studies. Some initial processing had already been performed, including base calling (identifying the exact DNA base at each given genotyped location) and population stratification (a method to identify hidden influences from ancestral background which may otherwise confound results). As such, the current methods concern analyses subsequent to this step.

Methods

The datasets we used were generated as part of the follow-up fine-mapping project Large Lupus Association Study 2 (LLAS2) [338] based on the results from our lupus GWAS [142]. The data contained genotype information at hundreds of thousands of bases across the genome for tens of thousands of individuals. Additionally, there was also collected at the time of participant recruitment clinical information on a variety of lupus-related phenotypes.

Logistic Regression

The basic analysis upon which all subsequent tests are based consisted of logistic regression of the SNP genotypes- comparing cases and controls for increased occurrence of any given SNP in one group versus the other. This analysis was performed in PLINK (v1.07) [339]. We began with data cleaning steps, filtering out SNPs with more than 10% missing genotypes, then removing individuals missing more than 10% genotype information, then removing SNPs with minor allele frequencies (MAFs) below 1% and finally, removing markers that deviate significantly from Hardy-Weinberg equilibrium. The admixture values for each individual estimating their ancestral origins were included as covariates in the analysis. The results from this test were used for most subsequent analysis and provide a broad picture of the overall association in the region.

Bayesian analysis

As a complementary test using a Bayesian statistical approach, and as a method of validation, we calculated the Bayes factor for each SNP using SNPTEST [340-342]. We then calculated the posterior probability for each SNP, which provides an estimate of the fraction of the total regional association accounted for by each individual SNP. This strategy has been described in ref [343] For this analysis we again controlled for population stratification by using the admixture values for each individual as coviariates in the analysis.

Imputation

Following the initial logistic regression analysis, we then imputed the region against the 1000 Genomes Project composite imputation panel in IMPUTE2 [245, 340, 344]. Imputation is used to fill in the missing genotypes at untyped SNPs from the original fine-mapping dataset. By leveraging the more complete regional structural information from the 1000 Genomes samples the complete genotype information for the common variants in the region can be accurately estimated for the individuals in our study. Imputed genotypes are required to meet or exceed a probability threshold of 0.5, and information measure of >0.4. Again highlighting a common difficulty in these analyses, the data format used in PLINK is incompatible with IMPUTE2, so there is a step of data conversion both before and after imputation. Following imputation, the data is again cleaned in PLINK using the same steps to ensure that no abnormalities were introduced during imputation. After cleaning, the dataset is ready for a logistic regression analysis using the same tests as before.

Conditional analysis

The results of an association test are useful in highlighting an area of the genome most likely to contain a variant responsible for the detected association. The challenge is that most often multiple variants are associated with the phenotype in question. Conditional analysis, specifically step-wise logistic regression, in which genotype information at specific variants is included in the logistic regression testing as a covariate can be useful in understanding the contribution of that variant to the overall signal at the locus. Conditional analysis is especially useful in identifying the number of

independent associations in the region. In a region with only one true association, even though there may be many highly associated variants across a large area of the genome, step-wise regression controlling for a single, highly associated variant, will account for most of the signal in the region, increasing the p values of all tested variants. When the reduction in signal is universal across all variants, and the individual p values are increased to be no longer significant, we can conclude that there is most likely only one source of association in the region.

In other regions, there may be multiple causal variants underlying the association in the region. In this case, controlling for the genotype at a single variant will reduce the signal for a portion of the associated variants in the region, but not for others. Single variant conditioning can identify which variants are associated with each overall effect, then multiple variant testing can be used to determine which variants can adjust for the entire signal in the region.

Haplotype estimation and testing

Haplotypes are the specific alleles within a genetic region that are commonly inherited together. Haplotypes occur as a result of the linkage disequilibrium (LD) within a region. LD is a measure of the propensity of two alleles to be inherited together and is related closely to their genomic distance from one another. Variants that are closer together are less likely to be separated by crossover events. Two variants in high LD are more likely than two variants with low LD to be inherited together and be part of the same haplotype. Haplotype blocks can be constructed based on measuring the LD in

the region. The method we use was described by Gabriel et al [345] and constructs blocks based on 95% confidence bounds based on D prime (a measure of LD). A haplotype block is defined as 95% of the variants in the block are all in “strong LD” with one another.

We estimate haplotype blocks using the Gabriel method in either PLINK or Haploview, the later having the advantage of displaying an LD plot where blocks and regional LD patterns can be visualized. Following haplotype block estimation, we again perform logistic regression association, testing each haplotype instead of each individual variant.

Meta analysis

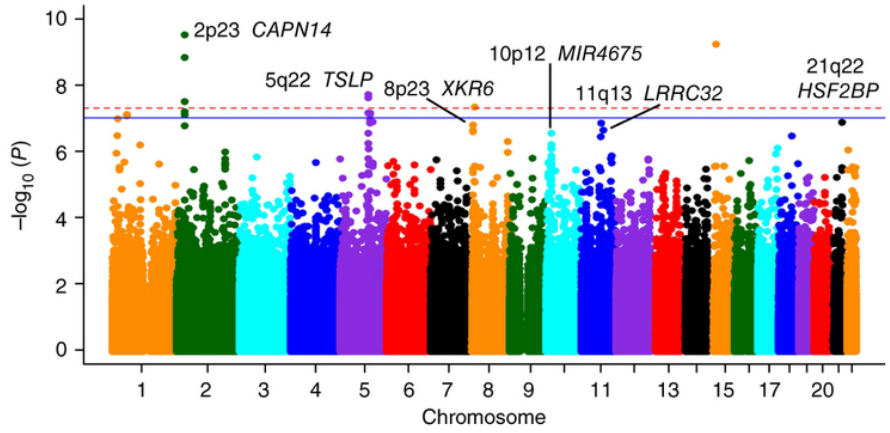
Meta analysis in genetic association studies is often used in the traditional manner to combine results from different studies. A related, more specific use of meta analysis, is to combine results from separate ancestral populations. Transancestral analyses in genetic association studies are a powerful way to identify shared disease risk and common risk variants amongst different ancestries [346]. Evolutionarily older populations, such as those originating in Africa, have had more time for crossover events and genetic recombination to occur, resulting in smaller haplotype blocks. This is in contrast to younger populations, such as those in Europe, where less recombination has occurred and haplotype blocks are larger. As a result of this difference between ancestral groups, meta analysis can often identify shared risk variants between the ancestries, leveraging the advantage of larger blocks in a younger population which

allow broad identification of associated regions, with the advantage of the smaller blocks of an older population which allow for some narrowing down of the association.

Data Visualization

A key aspect of our analysis pathway is data visualization. Visual inspection of results can be used to guide the next steps analytically, including highlighting regions of the genome for further refinement. Major computational errors in analysis can also be quickly detected by viewing the results graphically. We utilized the statistical software package R for our graphical output. R is a robust, open-source statistical graphing software package developed around the R programing language. We developed a series of scripts we use to prepare our association data for publication ready plots. For Manhattan plots (Figure 5), we used the R package qqman. For our association plots, we used ggplot and developed custom functions based on code provided by C. Lessard as seen in ref [347]. We have focused on presenting the data in such a way as to highlight imputed vs. directly genotyped

Figure 5: Manhattan plot of the P values obtained from the genome-wide association analysis.



Data are from 736 subjects with EoE and 9,246 controls over 1,468,075 genetic variants with MAFs greater than 1% in the subjects with EoE. The –log10 value of each probability is shown as a function of genomic position on the autosomes. Genome-wide significance (red dashed line; P ≤ 5 × 10–8) and suggestive significance (solid blue line; P ≤ 1 × 10–7) are indicated.

variants. The overall regional association is also highlighted as the LD between the top variant (or other select variants as needed) and the other variants in the locus (Figure 6). The Bayes factors identify the 95% credible set. The results of this script can be seen in our published work (Figure 10).

Figure 6: R script used to generate association plots.
One example of an R script used to visualize the results of a fine-mapping analysis including imputation and LD results. The recombination rate for the region is included and genes are plotted using information in biomart, a public database.

```
require(biomaRt)
#read in map and results
#This is the recombination map
combined3<-read.table("genetic_map_chr3_combined_b37.txt",header=T)
#These are the imputation results
chr3<-read.table("EA_apr.assoc.logistic",header=T)
#These are the results specific to LLAS2
llas2<-read.table("EA_pxk_admix_ci.assoc.logistic",header=T)
chr3Mb<-chr3$BP/1000000
llas2names<-llas2$SNP
combined3Mb<-combined3[,1]/1000000
#These are for plotting LD, if it's in the file
LD <- chr3$R2

#subset the file based on LLAS2 SNPs
sub<-chr3[chr3$SNP %in% llas2names,]
subMb<-sub$BP/1000000

#Get information for the SNP you'd like to highlight
rs76<-chr3[grep("rs4681677",chr3$SNP), ]
rs76Mb<-rs76$BP/1000000

#Draw lines to outline area of interest
#line1=58261741/1000000
#line2=58327540/1000000

#define genomic region for plots
rows<-nrow(chr3)
r_start<-chr3$BP[1]
r_end<-chr3$BP[rows]
r_startMb<-r_start/1000000
r_endMb<-r_end/1000000
row_p<-which.min(chr3$P)

#plot association
#Call quartz window and write to pdf. Name file as you like.
quartz(width=11,height=8,type="pdf",file="ea_pxk_imputed_not_new_rs46_may.pdf
")
par(mfrow = c(1, 1), mar=c(5, 5, 5, 5) + 0.1)
#plot- change xlim as needed- I manually set them for this plot.
```

```
plot(chr3Mb,-log10(chr3$P),xlab="Chr3 (Mb)",ylab=expression(paste(-
log[10],"(",italic(p),")")),xlim=c(58.15,58.55),ylim=c(0,11),cex=1.5,cex.lab=
1,cex.axis=1,lwd=1,las=1,pch=21,bg="royalblue")
title(main= "Imputed versus directly typed SNPs")
legend("right",c("imputed","direct","rs4681677"),pch=21,col="black",pt.bg=c("
royalblue","red","purple"))
#genome-wide significance line
abline(h=-log10(5E-8),lty=3,cex=10,col="darkred")
#Add points for LLAS2 on top of full data
points(subMb,-log10(sub$P),cex=1.5,pch=21,bg="red")
#add colored point highlighting SNP of interest
points(rs76Mb,-log10(rs76$P),cex=2,pch=21,bg="purple")

#Add more parameters for graph- these will be for recombination rate
par(new=T)
par(mfrow = c(1, 1), mar=c(5, 5, 5, 5) + 0.1)

#plot recombination rate
plot(combined3Mb,combined3$COMBINED_rate.cM.Mb.,type="l",xlim=c(58.15,58.55),
ylim=c(0,100),cex=1,cex.lab=1,cex.axis=1,lwd=1.5,xlab="",ylab="",xaxt="n",yax
t="n",col="blue")
axis(4,cex=1,cex.lab=1,cex.axis=1,lwd=1,las=1,col.axis="blue",col="blue")
mtext("Recombination Rate (cM/Mb)",side=4,line=2.5,cex=1,col="blue",crt=180)

#define build for biomart
mart<-useMart("ensembl",dataset="hsapiens_gene_ensembl")

#get list of genes in region
genes <-
getBM(attributes=c("ensembl_gene_id","start_position","end_position","exon_ch
rom_start","exon_chrom_end","strand","external_gene_id"),filters=c("chromosom
e_name","start","end"),values=list(3,r_start-
100,r_end+100),mart=mart,checkFilters=FALSE)

#plot genes

#split file by genes
gene_sets<- split(genes,genes$start_position)
gene_names <- c(names(gene_sets))

geneSymbol<- function (CCDS,file,label,dir,... ) {

      exonn<- nrow(CCDS)
      eStart<-CCDS[1,2]
      eEnd<- CCDS[exonn,3]
      hei<- 94
      eStart<-eStart/1000000
      eEnd<-eEnd/1000000
      wid<- 0.25
      rect(eStart,hei-wid,eEnd,hei+wid,col="dark blue",border=NA)


      if(dir==-1){
            text(x=(eStart+eEnd)/2,y=hei-4,label,col="dark
blue",font=3,cex=0.9)
            arrows((eStart+eEnd)/2+0.008,hei-1.5,x1=(eStart+eEnd)/2,y1=hei-
1.5,length=0.05,angle=20,col="dark blue")
      } else{
```

```
            text(x=(eStart+eEnd)/2,y=hei+4,label,col="dark
blue",font=3,cex=0.9)
            arrows((eStart+eEnd)/2-
0.008,hei+1.5,x1=(eStart+eEnd)/2,y1=hei+1.5,length=0.05,angle=20,col="dark
blue")
      }

}
#repeat for all genes in set
par(new=T)
for (i in 1:length(gene_sets)){
geneSymbol(CCDS=gene_sets[[i]],file=chr3,label=gene_sets[[i]][1,7],dir=gene_s
ets[[i]][1,6])
}
dev.off()
```

Bash scripts

In order to meet our goals of simplicity, usability and repeatability, we combine multiple steps of our analysis pipeline into bash scripts that can be run on the computational cluster. Bash scripts are small computer programs that feed instructions to the computational cluster on how many computational resources a given job will assigned, which programs will run, and which datasets will be analyzed. Figure 7 is an example of a bash script requesting 72 hours of computational time and 4 GB of memory. The programs to be loaded are PLINK, GTOOL and IMPUTE2. The first part of the program loops through a list of file names in a text file, performing the PLINK operations on each file sequentially as listed in the loop. Following the loop, one file is run through PLINK, selecting the region of interest. The file set is then converted to the format compatible with IMPUTE2 using GTOOL. Imputation is then run on the data in IMPUTE2. The files are then re-converted to a PLINK compatible format in GTOOL, and then the logistic regression analysis is performed in PLINK. At this point the data are ready for visualization and interpretation.

Figure 7: Bash script written for data cleaning, imputation, and association testing.

This script allows for the performance of multiple analytical steps form one submission to the computing cluster. Each individual step, including data cleaning, file conversions, imputation, and association analysis are contained within the script. This script allows the user to run imputation and association testing on a list of datasets simultaneously in one step, without needing to intervene at each step for each file.

```
#!/bin/bash
#BSUB -W 72:00
#BSUB -M 4000
#BSUB -e ./lsf_logs/%J.err
#BSUB -o ./lsf_logs/%J.out


#Load the required applications
module load plink
module load gtool
module load impute

fName="file_names.txt"

OFS=$IFS
IFS=$'\n'


#Data cleaning steps
for line in 'sed "$fName"'
do
name=`echo $line | cut -f1`
plink --noweb --bfile /data/"$name" --make-bed --geno 0.1 --out
/data/"${name}_g"
plink --noweb --bfile /data/"${name}_g" --make-bed --mind 0.1 --out
/data/"${name}_gm"
plink --noweb --bfile /data/"${name}_gm" --make-bed --maf 0.01 --out
/data/"${name}_mm"
plink --noweb --bfile /data/"${name}_mm" --make-bed --hwe 0.001 --out
/data/"${name}_mmh"
done

IFS=$OFS

#Convert the data into a compatible file type
plink --bfile /data/Omni1_mmh --recode --chr 1 --from-mb 58 --to-mb 59 --out
/data/OMNI1
gtool \
-P \
--family \
--ped /data/OMNI1.ped \
--map /data/OMNI1.map \
--og /data/OMNI1.gen \
--os /data/OMNI1.sample

#Impute the association results using the 1000 genomes data
impute2 -m /database/impute2_reference/genetic_map_chr1_combined_b37.txt \
-h /database/impute2_reference/ALL.chr1.haplotypes.gz \
-l /database/impute2_reference/ALL.chr1.legend.gz \
```

```
-g /data/OMNI1.gen \
-int 58e6 59e6 \
-buffer 20kb \
-iter 30 \
-burnin 10 \
-Ne 20000 \
-call_thresh 0.9 \
-o /data/OMNI1.impute2


#File conversion step
gtool \
-G \
--g /data/OMNI1.impute2 \
--s /data/OMNI1.sample \
--alleles \
--ped /data/OMNI1_imputed.ped \
--chr 3 \
--map /data/OMNI1_imputed.map \
--threshold 0.9 \
--sex sex \
--phenotype phenotype

#Reformat the file
awk '{OFS="\t"; print $1, $2, $3, $4}' /data/OMNI1_imputed.map >
/data/OMNI1_imputed2.map

mv /data/OMNI1_imputed2.map /data/$OMNI1_imputed.map

#Run the association test on the imputed data
plink \
--noweb \
--file /data/OMNI1_imputed \
--logistic \
--covar covar_all.txt \
--covar-name admixWA,admixIA,admixAS \
--out /data/OMNI1_imputed_log
```

Another advantage of using basic scripting language to run analysis is the ability to run the same analysis on multiple files simultaneously. One specific task that would have required weeks for one individual researcher to run is the conditional analysis described above. We have developed a strategy to identify the most important variants to the association signal by performing the step-wise regression analysis controlling for each variant in the region. As some regions after imputation can contain thousands of variants, this represents a significant time investment to run each test. We have utilized

a bash script to streamline this operation that will loop through each variant in the region one by one (Figure 8). The results are then compiled and written to a single file, allowing us to quickly identify the variants that control for the association signal to a given significance threshold.

Figure 8: Bash script written for looping through conditional analysis.
This script allows looping through and conditioning on each possible SNP in the dataset. This allows for an unbiased approach to the conditional analysis, where each variant is considered independently and with equal weight to all other variants. Each conditional analysis that results with residual association for a variant is written into a new file, allowing rapid assessment of which variants are able to account for all association in a region.

```
#BSUB -W 24:00
#BSUB -M 10000
#BSUB -e ./lsf_logs/%J.err
#BSUB -o ./lsf_logs/%J.out

#WARNING: run this program within a new, empty directory.
#Thousands of files will be generated.
#You need binary plink files and a SNPs.txt file to run this script
#The SNPs.txt file needs a list of all variants in the file.
#You can get this from your .bim or .map file.
#There can only be one SNP per line and not other information.

fName="~/data/SNPs.txt"

#Load required applications
module load plink
mkdir ~/data/out/adjusted_for_one_variant

#Adjust for every SNP in a binary plink dataset

for line in $(cat $fName)

do

#NOTE: It is necessary to change this code to reflect your genetic dataset.

  plink --noweb \
  --bfile ~/data/dataset \
  --logistic \
  --covar ~/data/covar.txt \
  --covar-name admixWA,admixIA,admixAS \
  --hide-covar \
  --condition "$line" \
  --out "~/data/out/${line}_conditioned"

#Identify all pairs of SNPs capable of removing all association
```

```
#p<0.01 in a .txt results file
#NOTE: The level of significance can be adjusted

  awk '$9<=0.01' "~/data/out/${line}_conditioned.assoc.logistic"
      > "~/data/out/${line}.lessthan0.01"

done

#move all lines from all files with variants that retain significance
#after adjustment to a new directory

mv *.lessthan0.01 ~/data/out/adjusted_for_one_variant

cd ~/data/out/adjusted_for_one_variant

#This file will contain a line for each file in the directory
#with the number of lines in each file.
#Each line contains a variant with residual association.
wc -l * > ~/data/out/results.txt
```

By adopting a consistent approach to our dataset analysis, we have the power to automate and standardize common tasks. With new datasets can we are able to significantly reduce total man-hours on any individual job as the basic tasks can all be run quickly and efficiently. In addition, standardization facilitates reproducibility and consistency.

Resources

PLINK- http://pngu.mgh.harvard.edu/~purcell/plink/

INPUTE2- https://mathgen.stats.ox.ac.uk/impute/impute_v2.html

R- http://www.r-project.org/

SNPTEST- https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html

METAL- http://www.sph.umich.edu/csg/abecasis/metal/

Haploview-http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview

Chapter 4: Lupus risk variants in the *PXK* locus alter B-cell receptor internalization

Samuel E Vaughn[1,2], Corinne Foley[3], Xiaoming Lu[2], Zubin Patel[1,2], Erin Zoller[2], Albert F Magnusen[2], Adrienne H Williams[4], Julie T Ziegler[4], Mary E Comeau[4], Miranda C Marion[4], Stuart B Glenn[5], Adam Adler[5], Nan Shen[1,6], Swapan Nath[5], Anne M Stevens[7,8], Barry I Freedman[9], Betty P Tsao[10], Chaim O Jacob[11], Diane L Kamen[12], Elizabeth E Brown[13,14], Gary S Gilkeson[12], Graciela S Alarcón[14], John D Reveille[15], Juan-Manuel Anaya[16], Judith A James[5,17], Kathy L Moser[5], Lindsey A Criswell[18], Luis M Vilá[19], Marta E Alarcón-Riquelme[5,20], Michelle Petri[21], R Hal Scofield[5,17,22], Robert P Kimberly[14], Rosalind Ramsey-Goldman[23], Young Bin Joo[24], Jeongim Choi[24], Sang-Cheol Bae[24], Susan A Boackle[25], Timothy J Vyse[26], Joel M Guthridge[5], Bahram Namjou[2], Patrick M Gaffney[5], Carl D Langefeld[4], Kenneth M Kaufman[2,27], Jennifer A Kelly[5], Isaac TW Harley[1,2,] John B Harley[2,27], and Leah C Kottyan[2,27,¶]

[¶]Corresponding author. (513) 636-1316, Fax: (513) 636-3668 E-mail: Leah.Kottyan@cchmc.org

1 Immunology Graduate Program and Medical Scientist Training Program, University of Cincinnati College of Medicine, Cincinnati, Ohio, United States of America

2 Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States of America

3 Spelman College, Atlanta, Georgia, United States of America

4 Center for Public Health Genomics and the Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America

5 Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, United States of America

6 Joint Molecular Rheumatology Laboratory of the Institute of Health Sciences and Shanghai Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai Institutes for Biological Sciences, and Chinese Academy of Sciences, Shanghai, China

7 Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, Washington, United States of America

8 Division of Rheumatology, Department of Pediatrics, University of Washington, Seattle, Washington, United States of America

9 Department of Internal Medicine, Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America

10 Division of Rheumatology, Department of Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, United States of America

11 Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America

12 Division of Rheumatology, Medical University of South Carolina, Charleston, South Carolina, United States of America

13 Department of Epidemiology, University of Alabama at Birmingham, Birmingham, Alabama, United States of America

14 Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, United States of America

15 Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, Texas, United States of America

16 Center for Autoimmune Disease Research, Universidad del Rosario, Bogota, Colombia

17 Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, United States of America

18 Rosalind Russell / Ephraim P Engleman Rheumatology Research Research Center, Department of Medicine, University of California San Francisco, San Francisco, California, United States of America

19 Division of Rheumatology, Department of Medicine, University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico, United States of America

20 Center for Genomics and Oncological Research, Pfizer-University of Granada-Junta de Andalucia, Granada, Spain

21 Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

22 United States Department of Veterans Affairs Medical Center, Oklahoma City, Oklahoma, United States of America

23 Division of Rheumatology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States of America

24 Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea

25 Division of Rheumatology, University of Colorado School of Medicine, Aurora, Colorado, United States of America

26 Divisions of Genetics and Molecular Medicine and Immunology, King's College London, London, United Kingdom

27 United States Department of Veterans Affairs Medical Center, Cincinnati, Ohio, United States of America

Abstract

Genome wide association studies have identified variants in *PXK* that confer risk for humoral autoimmune diseases, including systemic lupus erythematosus (SLE or lupus), rheumatoid arthritis and more recently systemic sclerosis. While PXK is involved in trafficking of epidermal growth factor Receptor (EGFR) in COS-7 cells, mechanisms linking PXK to lupus pathophysiology have remained undefined. In an effort to uncover the mechanism at this locus that increases lupus-risk, we undertook a fine-mapping analysis in a large multi-ancestral study of lupus patients and controls. We define a large (257kb) common haplotype marking a single causal variant that confers lupus risk detected only in European ancestral populations and spans the promoter through the 3' UTR of *PXK*. The strongest association was found at rs6445972 with $P < 4.62 \times 10^{-10}$, OR 0.81 (0.75 – 0.86). Using stepwise logistic regression analysis, we demonstrate that one signal drives the genetic association in the region. Bayesian analysis confirms our results, identifying a 95% credible set consisting of 172 variants spanning 202 kb.

Functionally, we found that PXK operates on the B-cell antigen receptor (BCR); we confirmed that PXK influenced the rate of BCR internalization. Furthermore, we demonstrate that individuals carrying the risk haplotype exhibited a decreased rate of BCR internalization, a process known to impact B cell survival and cell fate.

Taken together, these data define a new candidate mechanism for the genetic association of variants around *PXK* with lupus risk and highlight the regulation of

intracellular trafficking as a genetically regulated pathway mediating human autoimmunity.

Introduction

Systemic Lupus Erythematosus is the prototypical systemic autoimmune disease. Consequently, considerable effort has been devoted to understanding the genetic component of lupus risk. Now an understanding of causal mechanisms is approaching 10% of the 50 replicated genomic risk-loci that reach genome-wide significance [128].

Using a genome-wide association study (GWAS) design, we first identified association of a single nucleotide polymorphism (SNP) in the *PXK* locus with the occurrence of lupus in women of European descent [142]. The *PXK* locus association with lupus-risk has since been replicated in several studies [143, 185]. In these studies, the same variant identified in the initial GWAS was assessed in independent cohorts and the lupus-risk association was replicated [143, 185]. These studies did not include any biological or functional genomic follow-up of the replicated association of this variant [143, 185]. A fourth, recent study included a moderately powered fine mapping analysis of a European cohort and used expression quantitative trait locus analysis of nearby genes to argue for a role of *ABHD6* in the increased lupus risk. For the current analysis, we initiated a well-powered fine-mapping study aimed at identifying the likely causal variants and defining the biological mechanisms of lupus risk at this locus.

PXK is part of the sorting nexin (SNX) family of proteins, which are important for receptor internalization, organelle trafficking including endosomal trafficking, and other membrane-centric sorting functions. This is accomplished primarily through the PX-domain mediated binding of PI_3P [134, 348]. *PXK* was first identified and cloned by two independent groups in 2005 [132, 349]. Initial studies established that PXK is detectable

in most tissues with a primarily cytoplasmic distribution. A more recent study in COS-7 cells demonstrated that PXK co-localized with the EGFR. Furthermore, PXK facilitated EGFR internalization following ligand binding, which was found to be PX-domain dependent [141, 350]. *PXK* is widely expressed in the brain and blood [132], especially in B cells (Figure S1).

In this study, we identify a 257 kb region on chromosome 3, including all of *PXK* and a large region upstream of the gene that contains the lupus association signal. These results are confirmed via Bayesian analysis by which we identify a credible set consisting of 172 variants that explain 95% of the posterior probability in the region [343]. Through step-wise logistic regression analysis, we demonstrate that this region contains a single genetic effect.

Many studies in mice and humans highlight a central role of B cells in the etiology of lupus. Using the transcriptome of various B cell subsets, Hu *et al.* demonstrated that B cells and especially transitional B cells are enriched for transcripts near genetic variants associated with increased lupus risk [246]. Autoreactive B-cells play a critical role in the development of autoantibodies that lead to immune complex deposition and lupus-associated tissue damage [351]. In murine studies, mice without B cells are largely protected from lupus-like disease [352]. Furthermore, murine studies established a clear role for B cells in the dysregulated cytokine production and T cell activation associated with lupus-like autoimmunity [353, 354].

B cell signaling through the B cell receptor (BCR) plays a critical role in the development of autoimmunity in lupus [355, 356]. For example, BCR internalization facilitates Toll like receptor ligand internalization and signaling. Furthermore, type I IFN- a key cytokine known to play a role in lupus- is known to promote rapid B cell receptor internalization [357, 358]. Based on its role in regulating EGFR, we hypothesized that PXK participates in BCR internalization and lupus-associated variants in the *PXK* locus differentially regulate BCR internalization. Indeed, we demonstrate that PXK colocalizes with the BCR and that the risk variants are associated with a decrease in BCR internalization. Knockdown of PXK replicates this phenotype, confirming the direct involvement of PXK in BCR trafficking.

Methods

Subjects and Study Design

We used a large collection of samples from case-control subjects from multiple ethnic groups. These samples were from the collaborative Large Lupus Association Study 2 (LLAS2) [338] and were contributed by participating institutions in the United States, Asia, and Europe. According to genetic ancestry, subjects were grouped into ethnic groups including European American (EU), African American (AA), Asian and Asian American (AS), and Hispanic American (HA). Informed consent was obtained from all subjects using Institutional Regulatory Board approved consent documents. All lupus patients met the American College of Rheumatology (ACR) criteria for the classification of lupus [24].

Genotyping of genetic variants and Sample Quality Control

We genotyped 58 single-nucleotide polymorphisms (SNPs) covering the entire *PXK* region (58.1-58.5 MB on Chr 3, Build 37), as part of a larger custom genotyping study. The variants were chosen based upon the results of a candidate association study of 720 women of European ancestry and 2,337 controls [142]. Specifically, the variants were chosen to span the association interval identified with the Infinium HumanHap330 array. Genotyping of SNPs was completed with Infinium chemistry on an Illumina iSelect custom array according to the manufacturer's protocol. The following quality-control procedures were implemented to identify SNPs for analysis: well-defined clusters for genotype calling, call rate >90% across all samples genotyped, minor allele frequency (MAF) >1% (except for the rare variant analysis as described below), and $p>0.05$ for differential missingness between cases and controls (the total proportion missing was <5%). One marker with evidence of a departure from Hardy-Weinberg proportion expectation ($p<0.0001$ in controls) was removed from the initial analysis.

We removed individuals with a call rate <90% or excess heterozygosity. The remaining individuals were examined for excessive allele sharing as estimated by identity-by-descent (IBD). In sample pairs with excessive relatedness (IBD>0.4), one individual was removed from the analysis on the basis of the following criteria: (1) remove the sample with the lower call rate, (2) remove the control and retain the case, (3) remove the male sample before the female sample, (4) remove the younger control before the older

control, and (5) in a situation with two cases, remove the case with the less complete phenotype data available. Discrepancies between self-reported and genetically determined gender were evaluated.

Ascertainment of Population Stratification

Genetic outliers from each ethnic and/or racial group were removed from further analysis as determined by principal component (PC) analysis and admixture estimates (Figure 1 of ref [241] and ref [131, 359]). We used 347 ancestral informative markers (AIMs) from the same custom genotyping study that passed quality control in both EIGENSTRAT [131] and ADMIXMAP [130, 360] to distinguish the four continental ancestral populations, allowing identification of the substructure within the sample set [129, 361]. The AIMs were selected to distinguish four continental ancestral populations: Africans, Europeans, American Indians, and East Asians. We utilized principal components from EIGENSTRAT outputs to identify outliers of each of the first three PCs for the individual population clusters with visual inspection.

Statistical Analysis - Workflow

The analysis was initiated by assessing the association of genotyped variants in each of the four ancestral cohorts individually as done previously at another locus [362]. Strategically, we analyzed the genotyped, then imputed variants, performed full haplotype analysis, executed linkage disequilibrium analysis, and finally built a statistical model to account for the lupus-associated variability in European ancestry. In building

the model of association in European Americans, we comparatively evaluated every variant in the region for its ability to better account for the lupus-associated genetic variation.

Statistical Analysis - Frequentist approach

We tested each genetic variant for association with lupus using logistic regression models (frequentist approach) that included three admixture proportion estimates as covariates as implemented in PLINK v 1.07 [339, 340]. The additive genetic model is the primary model of inheritance. Other models were subsequently considered, but only if substantially superior. We performed a Cochran-Armitage trend test, a genotypic test, and tested both the dominant and recessive gene models. Logistic regression using an additive model remained the best model.

Step-wise logistic regression was performed to identify those genetic variants independently associated with the development of lupus in PLINK. For these analyses, the allelic dosage of a specific variant was added to the logistic model as covariates in addition to the admixture estimates. Haplotypic associations were assessed using logistic regression incorporating admixture measurements as covariates.

Linkage disequilibrium (LD) and haplotypes were determined with PLINK and HAPLOVIEW v 4.2 [363-365]. We calculated haplotype blocks for those haplotypes present at >3% frequency using the 4 gamete rule algorithms with a minimum r^2 value of 0.8. Haplotypic associations were performed in PLINK using a 200 kb sliding window approach.

Statistical Analysis - Bayesian approach

Using SNPTEST, we calculated the Bayes factor (BF) for each genetic variant: the probability of the genotype configuration at that genetic variant in cases and controls under the alternative hypothesis that the variant is associated with disease status divided by the probability of the genotype configuration at that variant in cases and controls under the null hypothesis that disease status is independent of genotype at that variant as previously described (we used the methods developed and introduced in ref [343]). We used three admixture estimates as covariates, as we did for the frequentist approach. Large values of the Bayes factor (BF) correlate to robust evidence for association, as small p-values correlate to strong evidence in a frequentist approach. For our well-powered study, the Bayes factors (BFs) of the variants were highly correlated with the p-values (consistent with ref [366]). We used the additive model. The linear predictor is $\log(p_i/(1-p_i)) = \mu + ßG_i$, and the prior is $\mu \sim N(0,1^2)$, $ß \sim N(0,0.2^2)$ (variables are defined in the supplementary note in ref [343]).

To identify the variants most likely to be driving the statistical association we calculated a posterior probability under the assumption that any of the variants within a single genetic effect could be causal and that only one of these variants is causal for each genetic effect. Variants with a low posterior probability are highly unlikely to be causal regardless of the allele frequency or presence of the actual causal variant in the analysis, following the procedure as presented [343].

Re-sequencing

We re-sequenced the *PXK* region as described previously [347]. DNA from European American subjects included in the current genotyping experiment was sequenced. To assess the accuracy of sequence-based SNP calling, we cross-referenced the sequenced and genotyped allele calls. Briefly, 3–5 micrograms of whole genomic DNA from each sample was sheared and prepared for sequencing with an Illumina Paired-End Genomic DNA Sample Prep Kit. Targeted regions of interest from each sample were then enriched with a SureSelect Target Enrichment System utilizing a custom-designed bait pool (Agilent Technologies). Post-sequence data were processed with Pipeline software v.1.7 (Illumina). All samples were sequenced to minimum average fold coverage of 253. Variant detection and quality control were also performed as previously described [347].

Imputation to composite 1000 Genomes reference panel

To detect associated variants that were not directly genotyped, we imputed the *PXK* region with IMPUTE2 and using a composite imputation reference panel based on 1,000 Genomes Project sequence data freezes from December 2013 [245, 340, 344]. Imputed genotypes were required to meet or exceed a probability threshold of 0.5, and information measure of >0.4, and the same quality-control criteria threshold described for the genotyped markers. In the statistical analyses, the probability threshold from each imputed value was incorporated into the statistical analysis using SNPTEST. The overall genotype-imputed variant concordance rate was >93%.

Rare variant analysis

Rare variant analysis was performed in SVS Golden Helix SNP & Variation Suite v7.6.10 on the re-sequenced dataset and the imputed dataset. Rare variants were filtered based upon call rate and minor allele frequency <0.01. The one-sided kernel based adaptive cluster algorithm was performed using the hyper geometric kernel type with 1000 permutations in SVS GoldenHelix. For these studies, the initial analysis was performed using sequencing data from 92 lupus cases and 114 controls. Because this initial study was underpowered to identify an increased rare variant burden, we extended this analysis to an imputed dataset of all 4220 European lupus cases and 3803 European lupus controls. The accuracy of rare variant imputation, while limited, has been previously validated [367, 368].

Power analysis

Power analysis was performed using the Genetic Power Calculator [369].

Cell culture

Lymphoblastoid cell lines (LCLs) derived from control study participants without lupus were generated as previously described [338] and maintained at 37° in RPMI 1640 media containing 10% heat-inactivated fetal bovine serum (FBS) and anti-biotic/anti-micotic (Gibco/Life Technologies, Grand Island, NY). Peripheral blood mononucelar cells (PBMCs) were isolated using Ficoll-gradient centrifugation in SepMate tubes

(Stemcell Technologies, Vancouver, BC, Canada) from fresh blood collected in acid citrate dextrose (ACD) tubes with informed consent from healthy donors. B cells were isolated using the B-cell isolation kit-II (Miltenyi Biotech, Auburn, CA).

RNA purification and expression analysis

Total RNA was purified from LCLs using the RNeasy kit (Qiagen Valencia, CA). RNA was reverse-transcribed with the High-Capacity RNA-to-cDNA kit (Applied Biosystems, Grand Island, NY). Gene expression was determined by real-time PCR using a 7500 Real-time PCR system or Viia 7 system (Applied Biosystems). Relative quantification was calculated by the comparative CT method [370]. Briefly, expression levels for each target were normalized to levels of 18S in the same well. All samples were then normalized to a selected sample and the data was exponentially transformed.

Protein expression was quantified in LCLs by flow cytometry. Cells were permeabilized and stained using reagents from Beckton, Dickinson and Company (BD, Franklin Lakes, NJ). Following staining, fluorescence was determined on an LSRFortessa analyzer (BD). The geometric median fluorescence intensity (gMFI) was calculated in FlowJo (Tree Star, Ashland, OR) and used as an estimate of protein expression. All experiments were run in triplicate. Statistical analysis and data visualization were done with GraphPad Prism (GraphPad Sftward, La Jolla, CA).

BCR internalization and colocalization

BCR internalization was adopted from the method described by Malhotra et al [371, 372]. Briefly, LCLs or freshly isolated PBMCs from subjects without lupus were incubated with anti-BCR for 30 min at 4° Celsius. Cells were then placed at 37°C for the indicated times. Cross-linked BCR remaining on the surface was then detected with fluorescently labeled secondary antibody with the LSRFortessa Cell Analyzer. Percent internalization was obtained by measuring the GMFI at each time point and then calculating the change as follows: (gMFI no internalization (4°. sample) – gMFI internalization (37°. sample)) ÷ (gMFI no internalization (4°. sample)) x 100. Statistical analysis was performed in R using the packages “nlme” and “lmerTest”. Cells were imaged with a confocal microscope and the ImageStreamX (Amnis, Seattle, WA) for co-localization analysis. All experiments were run in triplicate.

shRNA transfections

shRNA were obtained from the CCHMC Lenti-shRNA Library Core, utilizing the Sigma Mission system. Viral vector DNA was isolated using Endofree Plasmid Maxi Prep kits (Qiagen). Lentiviral transduction was performed by the CCHMC Viral Vector Core. Viral particles containing the shRNAs were used to transfect LCLs from control subjects. Cells were plated onto retronectin-coated plates, viral particle supernatant added, and spun for 45 min at 1300 x G at 22°C. After resting at 37°C overnight, spinfection was repeated. Cells were transferred to selection media containing 2 ug/ml puromycin after 48 hrs. After non-transfected cells began to die (approximately 3-5 days) cells were tested for *PXK* expression and maintained for use in subsequent experiments.

Materials

Anti-BCR ($F(ab')_2$ antibodies to IgG/IgM and Alexa Fluor 647-conjugated donkey anti-human IgG/IgM) were from Jackson ImmunoResearch (West Grove, PA). Anti-PXK antibody was purchased from Abcam (Cambridge, MA). Anti-LAMP1 antibody was purchased from BioLegend (San Diego, CA). Secondary antibodies were purchased from Life Technologies (Carlsbad, CA). Retronectin was purchased from Clonetech (Mountain View, CA). Puromycin was purchased from Invivogen (San Diego, CA). All other antibiotics and cell culture reagents were purchased from HyClone (Logan, UT) and Gibco (Life Technologies). Quantitative PCR was performed using TaqMan probes from Life Technologies.

Results

We genotyped 57 useful markers from the *PXK* locus in a transancestral group of 18,286 lupus cases and controls. Our study included 8,023 individuals of European descent (EA), 3,740 individuals of African American descent (AA), 2,481 individuals of Hispanic American descent (HA) and 2,652 individuals of Asian descent (AS). To better capture the total variation in the locus, we imputed against a composite reference panel derived from the 1000 Genomes Project [373] for a final dataset of 269-835 variant markers, depending on ancestry, all with a minor allele frequency (MAF) greater than 0.01 (Table 5).

Population	Case	Control	Male	Female	Genotyped	Imputed (MAF > 0.01)
European (EA)	4220	3803	1696	6327	57	477
African American (AA)	1719	2021	766	2974	57	835
Hispanic (HA)	1599	882	244	2237	57	593
Asian (AS)	1310	1342	301	2351	57	269

Cleaning steps are described in methods.

Table 5. Summary statistics for each study population. Cleaning steps are described in methods.

We performed logistic regression analysis to reveal a large association region of 257 kb extending from just upstream of *PXK* beginning in *ABHD6* and extending through *RPP14* in individuals of European descent (Figure 9a). Overall, the most significant SNP was rs6445972 with $P < 4.62 \times 10^{-10}$, OR 0.81 (0.75 – 0.86). The most significant, directly genotyped SNP was rs4681677 with $P < 2.00 \times 10^{-9}$, OR 0.81 (0.76 – 0.87). Because rs4681677 was genotyped and not assessed solely through imputation, we used this SNP to classify cell lines as risk or protective for subsequent biological experiments. We found no evidence of the lupus-association at this locus in the other populations studied (Figure 9b). rs6445972 has an allele frequency of 30.0% in the EA cohort, 10.4% in the AA cohort, 22.9% in the HA cohort, and 0.2% in the AS cohort, suggesting that a largely European-derived allele is driving the very modest statistical association in the cohorts with European admixture.

Analysis of linkage disequilibrium (LD) in the region revealed high LD between the most strongly associated variants (Figure 10a). No single haplotype (using either blocks constructed from continuous groups of variants or the most highly associated variants) outperformed the single variant association model

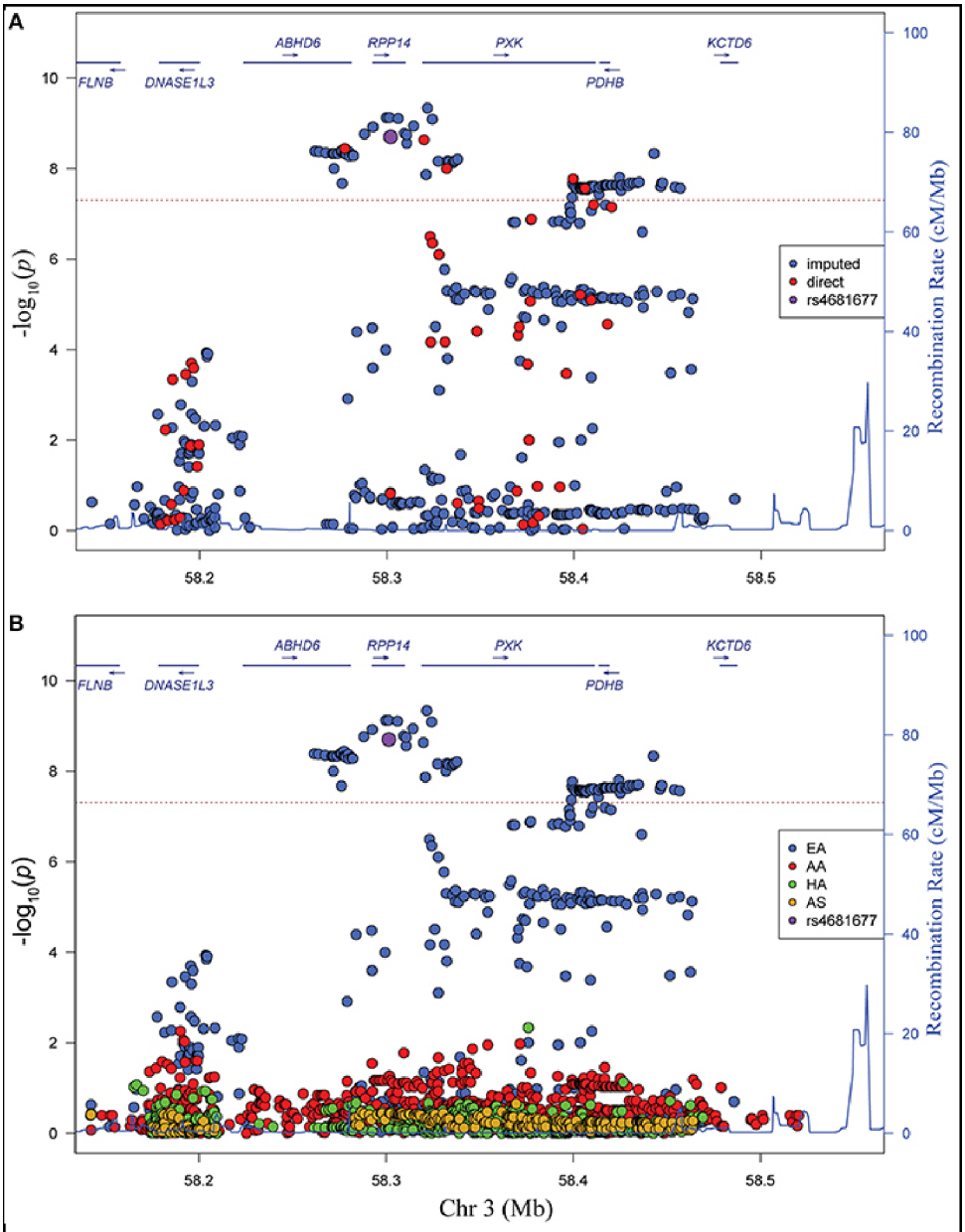


Figure 9. Variants in the PXK locus are associated with lupus. **(A)** Plot of –log10(p) values for directly genotyped (red) and imputed (blue) SNPs spanning the previously identified region of association with lupus. Gene locations are indicated above with arrows indicating direction of transcription. **(B)** Plot of –log10(p) values for different ancestral cohorts. EA, European ancestry; AA, African American ancestry; HA, Hispanic and Amerindian ancestry; AS, Asian ancestry. While not the most highly associated SNP, rs4681677 (purple) was used to choose samples for biological studies because it was genotyped; this SNP is in high LD (r2 > 0.9) with the imputed SNPs that are more highly associated.

(data not shown), supporting the conclusion that the association in the region is due to a single genetic variant. In a complementary Bayesian analysis, we find a similar pattern of association, consistent with results from our frequentist logistic regression analysis. The 95% credible set [343] contains 172 variants spanning a 202 kb region (chr 3: 58261741-58463411) (Figure 10b).

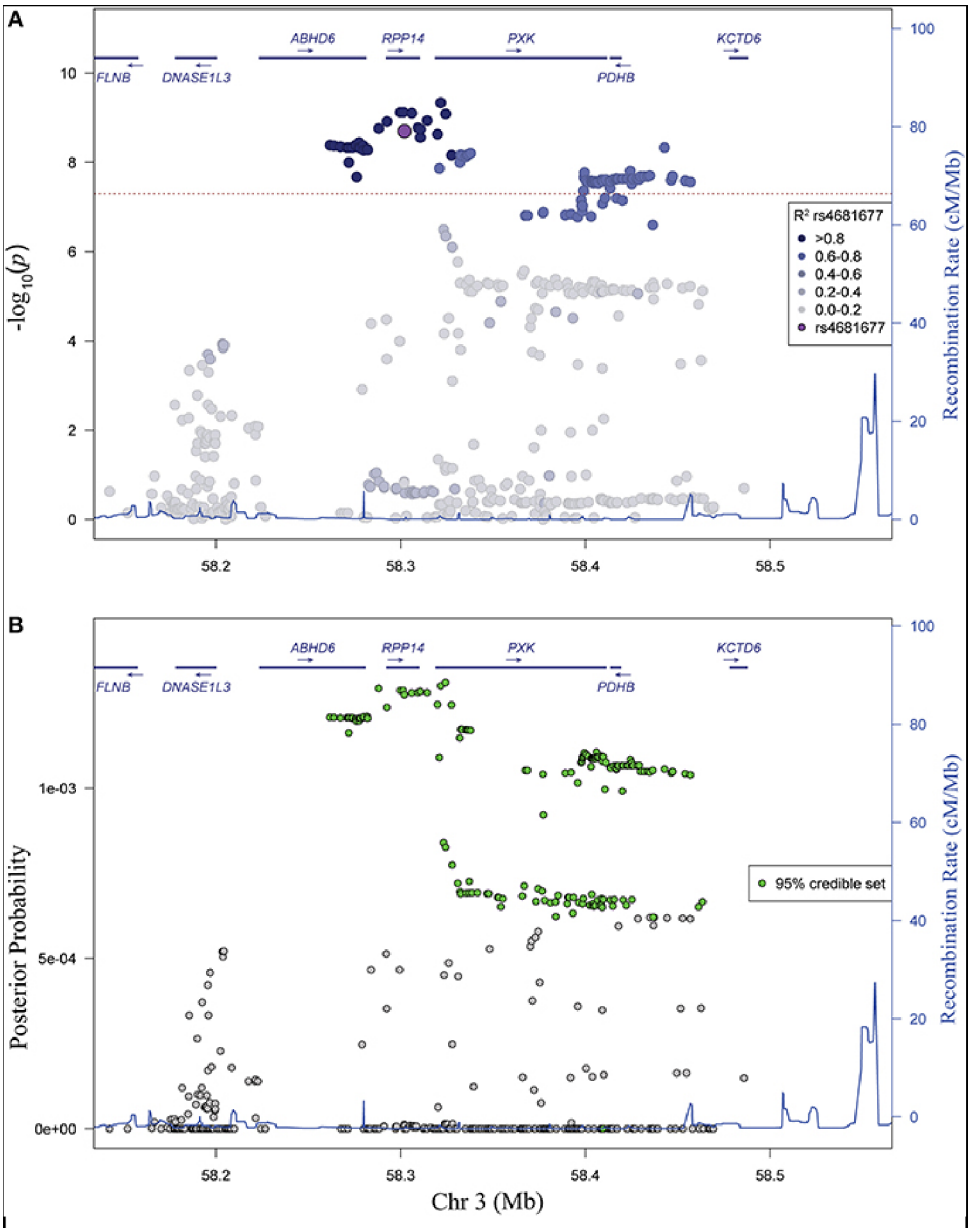


Figure 10. Evaluation of linkage disequilibrium in the PXK locus and Bayesian association analysis. (A) Association plot of genotyped and imputed SNPs shaded to identify the significant LD in the region. Colors correspond to r2-values, with darker shade indicating higher r2-values. (B) Bayesian analysis of variants in the PXK locus. Plot of posterior probability calculated for each variant in the locus. Variants included in the 95% credible set based on posterior probabilities are colored green.

To identify candidate variants in the EA population, we performed stepwise logistic regression to evaluate the ability of variants within the associated haplotype to explain all of the lupus-associated variation. Adjusting for any one of the top variants in the region eliminated the association signal, supporting the model that there is only one association in the region (Figure 11). We sequentially tested each variant in our dataset individually in our conditional analysis in an attempt to isolate groups of variants that may disproportionately carry the association signal in the region. While we did find clear groups, we found those variants that were able to adjust for the largest portion of the lupus-associated variation were in high LD (r^2>0.8) with the most strongly associated variants. Notably, the SNP identified in our original GWAS study, rs6445975, did not account for lupus risk as completely as the most strongly associated variants in the current study (Figure 11 and Figure S2).

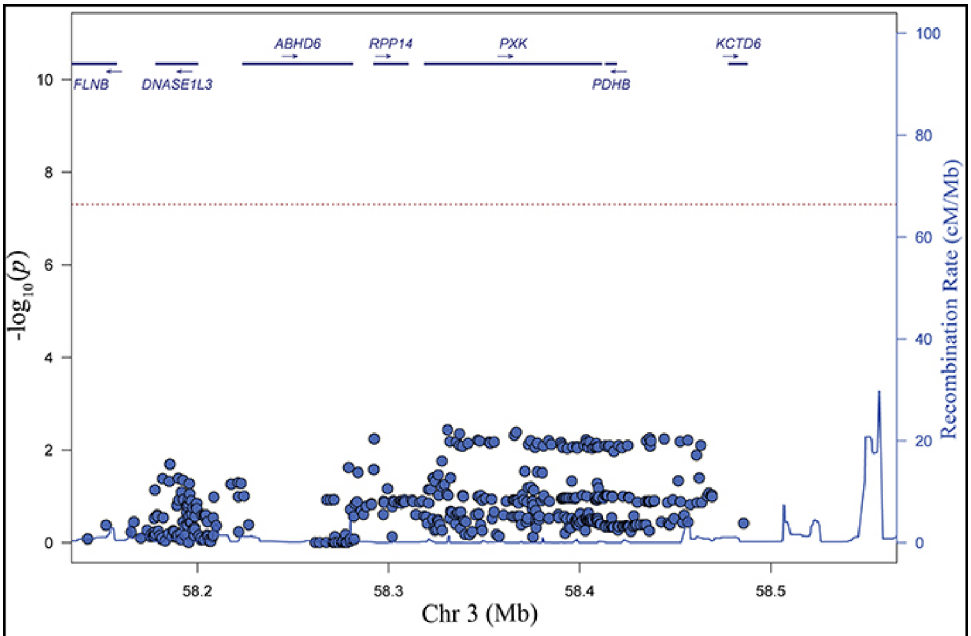


Figure 11. Step-wise logistic regression analysis reveals one independent genetic effect at the PXK locus. Plot of P values (–log10) following adjustment for genotype at rs4681677. Adjustment for other SNPs with r2 > 0.9 with rs4681677 can also account for the genetic association (P < 0.01) in the region.

It remained possible that rare variants were driving the lupus association at the *PXK* locus. To test this possibility, we performed deep sequencing of the region in 92 cases and 114 controls of European ancestry. We found no statistical association of any of the variants with frequencies less than one percent using a logistic regression analysis with an additive model; furthermore, we found no increased burden of rare variants in any of the genes in the region in the cases compared to the controls (data not shown). Despite the limitations of imputed rare variants, we repeated the rare variant burden test on the complete European population and did not find evidence of increased rare variants in the lupus cases.

In order to assess the hypothesis that lupus-associated variants were affecting the gene expression, we measured expression levels of all 5 genes in transformed B cell lines from control subjects and found no difference between cell lines with homozygous risk and homozygous non-risk genotypes (Figure S3). We

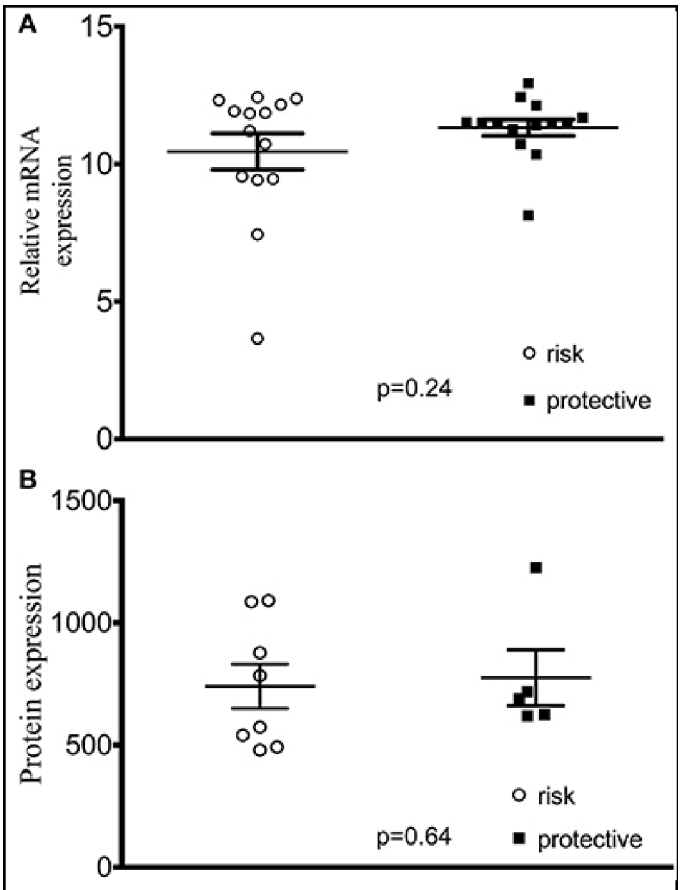


Figure 12. PXK expression in cell lines derived from study participants. **(A)** Plot of ΔΔCt values of PXK mRNA expression using TaqMan probes. (n = 14 risk/14 protective) **(B)** PXK expression at the protein level using flow cytometetry. (n = 8 homozygous risk/5 homozygous protective). Data for both A and B are representative of 3 independent experiments using control cell lines. Statistical analysis using a Student's t-test was performed in PRISM.

did not find any difference in mRNA or total PXK protein expression (Figure 12A-B).

PXK is highly expressed in the B cell lineage, with the highest expression in mice observed in both transitional and follicular splenic B cells [246]. The transitional B cell was recently implicated as one of the cell types critical to lupus pathogenesis based on a combined genetic and cell-specific expression analysis [374]. Given the expression of *PXK* in B cells, the fact that PXK has been shown to play a role regulating cell surface receptor expression, and the importance of B cells to the pathology of lupus, we hypothesized that PXK participates in the internalization of the BCR. We first evaluated colocalization between PXK and the BCR in B cells following BCR crosslinking on the cell surface. We found moderate steady-state colocalization between PXK and the BCR at baseline that increased following BCR crosslinking and continued to increase with cellular internalization (Figure 13). We then tested the hypothesis that the lupus-associated risk variants at the *PXK* locus affected the rate of BCR internalization by measuring the internalization of the BCR in cells derived from

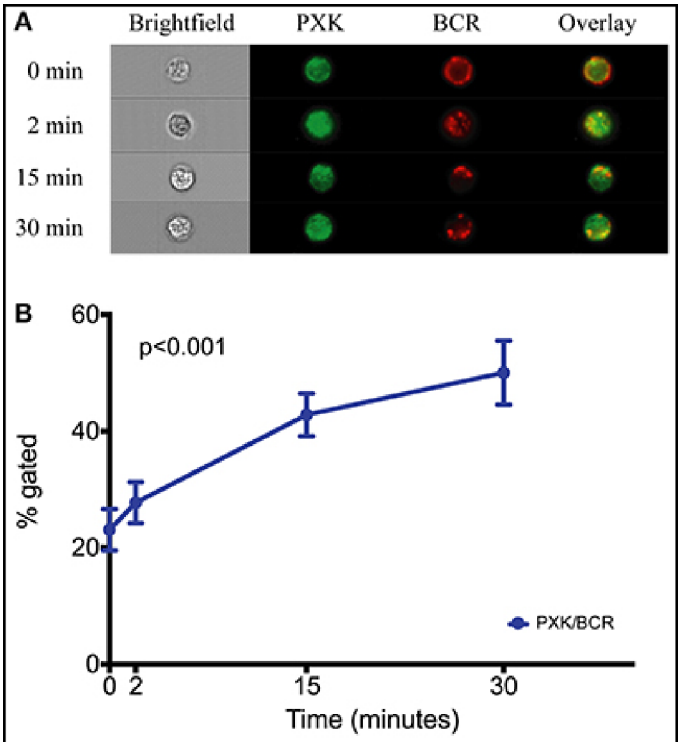


Figure 13. PXK colocalizes with the BCR following receptor crosslinking. **(A)** Quantification of colocalization between PXK and the BCR following BCR crosslinking in primary B cells from subjects without lupus. **(B)** Representative images showing colocalization. Statistical analysis using a One-Way ANOVA was performed in PRISM.

individuals with known genotypes. Following BCR crosslinking, we measured receptor internalization with flow cytometry and found that cells carrying the homozygous risk genotype displayed a decrease in the amount of BCR internalization compared to cells homozygous for the protective genotype (Figure 14). This finding was consistent across all the measurements over the time course of the experiment.

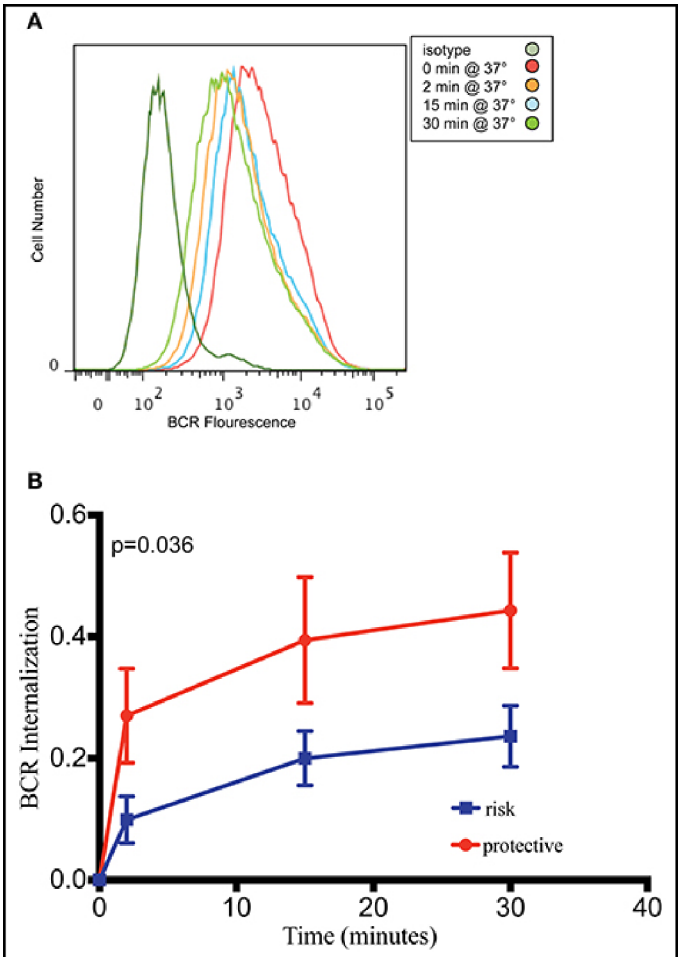


Figure 14. B cells carrying PXK risk allele demonstrate decreased internalization of the BCR. **(A)** Representative flow cytometry histogram showing decreasing fluorescence following internalization of the BCR from the B cell surface using patient-derived LCLs. **(B)** Quantification of BCR internalization as measured by the percent internalization (decrease in gMFI) following receptor crosslinking. Results of linear mixed model analysis as performed in R are shown. This experiment was performed using 20 cell lines from subjects without lupus in three independent experiments. The results of these independent experiments were combined and presented in **(B).**

To assess the role of PXK in the BCR internalization phenotype, we decreased *PXK* expression using shRNA by transfecting 5 study-derived cell lines from patient without lupus with shRNA targeted against *PXK* or scrambled controls (Figure 15A). Knockdown of PXK resulted in reduced internalization of the BCR, especially at the early time points (Figure 15B), and this reduction was directly correlated with *PXK* expression (Figure 15C).

Discussion

We performed a fine-mapping study to

define the genetic variants and the gene most likely to be causal for increased lupus risk. We genotyped 18,286 cases and controls from four ancestral populations, leveraging imputation to increase our coverage of the target locus. We used both logistic regression,and Bayesian methods to verify our findings that the disease risk from this locus is due to highly correlated variants in a large region including *PXK*. Additionally, we complemented our genetic fine-mapping with biological experiments, identifying a genotypic change in B cell receptor internalization that is associated with the presence of lupus-risk variants.

Analysis of the *PXK* locus in non-European populations revealed no significant association in the region (Figure 9B). The allele frequency of the

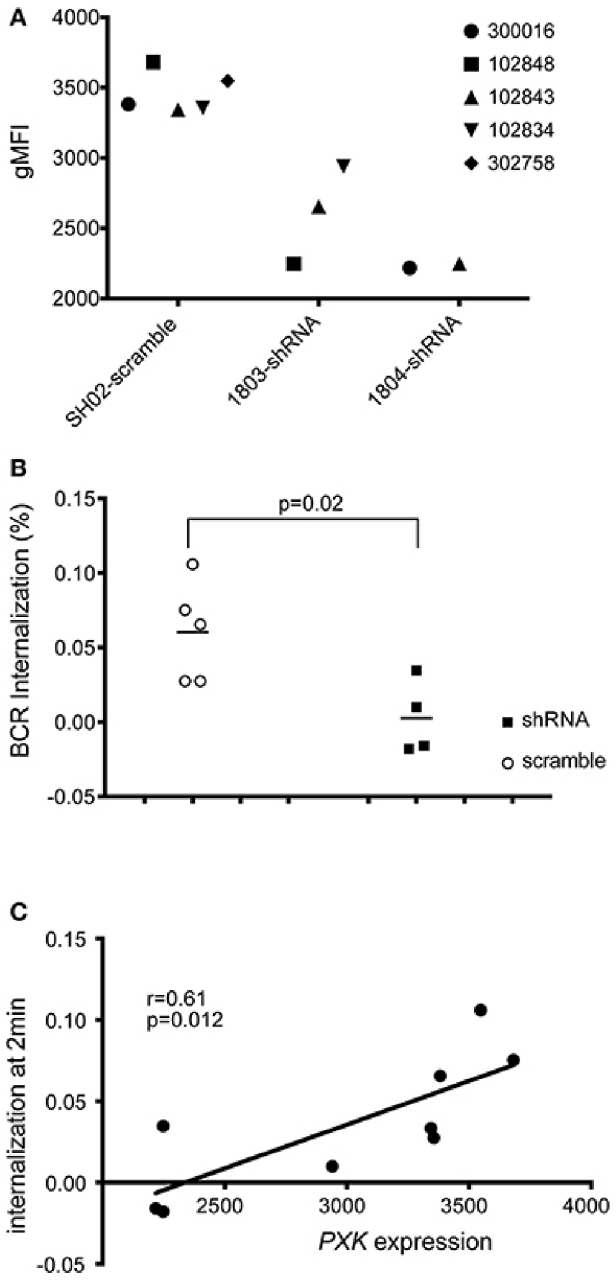


Figure 15. shRNA knockdown of PXK disrupts BCR colocalization. **(A)** PXK protein expression following shRNA knockdown in 5 LCL lines from non-lupus subjects (expression was assessed in triplicate for each cell line). **(B)** Summary of BCR internalization at 2 min in LCLs following viral transduction with either an shRNA against PXK or a scrambled control. Results from unpaired Student's t-test shown. **(C)** Correlation between PXK protein expression and BCR internalization following shRNA knockdown. The significance was assessed using a Pearson correlation. The graphs in **(B,C)** are representative of the results of three independent experiments.

most highly associated variants differed amongst the ancestral groups, and given the observed allele frequencies, power analysis indicates that a much larger non-European population will be needed to detect a similar association of this magnitude at this locus (data not shown). Thus the lack of an association in the other populations may be due entirely to the difference in power.

We found a 50% decrease in BCR internalization kinetics between risk and non-risk genotypes. This change in BCR internalization may have a small effect in the regulation of B cell signaling; however, it could also be meaningful in the context of other additive changes to the B cell signaling pathway [128]. There are several mechanisms through which this functional change could affect the risk of autoimmunity. For example, stimuli that would be sufficient to develop strong B cell activation with subsequent negative selection may now be less likely to result in elimination due to attenuated BCR internalization. Indeed, the persistence of auto-reactive clones in the periphery may play an important role in the context of lupus pathogenesis.

PXK colocalized with the BCR and affected internalization, indicating a potential role for PXK in the regulation of BCR signaling. Importantly, the colocalization increased as the BCR is internalized into the cell, suggesting a role for PXK beyond the cell surface. Furthermore, we find that cells carrying the lupus-risk haplotype have a decreased amount of BCR internalization when compared to cell lines carrying the non-risk

haplotype, an allelic functional change that we present as a candidate causal mechanism of increased lupus risk at this locus.

When *PXK* is specifically knocked down using shRNA, BCR internalization is decreased, confirming a clear role for PXK in regulating BCR internalization. These results support the conclusion that the allele-specific changes in BCR internalization we detect are most likely attributable to PXK. We do not yet understand the genetic mechanism behind this alteration in BCR internalization. We were unable to detect a difference in overall *PXK* expression. It may be that some other molecular differences are occurring after *PXK* expression, such as alternative splicing or differential posttranslational modification. Future studies will be directed at detecting these changes. While *ABHD6* eQTLs were shown to correlate with lupus-association, there were also *PXK* eQTLs that were associated with lupus [149]. The *PXK* locus remains complicated and future work will be important to continue to unravel the specific genetic variations underlying the lupus association.

We and others refer to this locus as the "*PXK* locus" [142, 149, 185, 298, 374]. Although our step-wise logistic regression analysis of common variants (Figure 11) makes it unlikely, it is still possible that there is a contribution of rare variant(s) that we missed. Of the 5 genes in the region, *PXK* and *DNase1l3* are the only genes with appreciable expression in immune cells based on evaluation of public databases (data not shown). *DNASE1L3* encodes the protein DNase 1-like 3, and is the only gene that has been

investigated functionally in the context of lupus. A loss-of-function frameshift mutation in *DNASE1L3* was found to be associated with early-onset lupus and lack of detectable *DNASE1L3* transcripts in an autosomal recessive manner in six families of Arab descent [375]. *DNASE1L3* was also recently identified as a risk gene in systemic sclerosis [376]. The variant identified in that study shows no association with lupus in our dataset (p=0.3067). *DNASE1L3* is removed from the peak association in the frequentist, logistic regression analysis and no members of the 95% credible set in the Bayesian analysis were located in or within 25kb of *DNASE1L3*. It remains possible that variants many hundreds of killobases away from the promoter of a gene could affect that genes transcription [377]; however, while both of our groups found detectable *DNASE1L3* transcripts, neither our group nor other groups have identified allelic expression of *DNASE1L3* in our cell lines [149]. Of the remaining genes, *ABHD6* has recently been suggested as the causal gene responsible for the association with lupus at this locus [149]. Mutations in *ABHD6* have been associated with multiple non-autoimmune phenotypes [378-382], but beyond *ABHD6* being in the associated region there is no evidence, at present, for ABHD6 in lupus pathogenesis.

Overall, we identified a limited haplotype of highly associated variants in the promoter and first exon of *PXK* that account for all of the lupus-association in the region. We performed biochemical analysis to demonstrate that PXK co-localizes with the BCR and affects BCR internalization. Furthermore, we identified an allelic decrease in PXK-BCR co-localization and BCR internalization in subjects expressing the lupus-risk haplotype.

Taken together, our work supports a model in which PXK increases lupus risk through the regulation of BCR internalization.

Acknowledgements

This work was supported by the Korea Healthcare Technology R&D Project, Ministry for Health and Welfare, Republic of Korea [HI13C2124 to SCB], the Alliance for Lupus Research [to BPT], the US Department of Veterans Affairs [IMMA 9 to JBH], the US Department of Defense [PR094002 to JBH] and the National Institutes of Health [R37AI024717 and P01AI083194 to JBH, RO1AR043814 and R21AR065626 to BPT, P01AR049084 to RPK, AR43727 to MP, P60AR062755 and UL1RR029882 to GSG and DLK, and K24AI078004, R21AI070304, UL1TR000154 and P30CA046934 to SMB, T32GM63483 to SEV and ZP. We would like to thank Summer Frank and Mai Li Zhu for their assistance in genotyping, quality control analyses, and clinical data management.

Conflict of Interest Statement

This research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

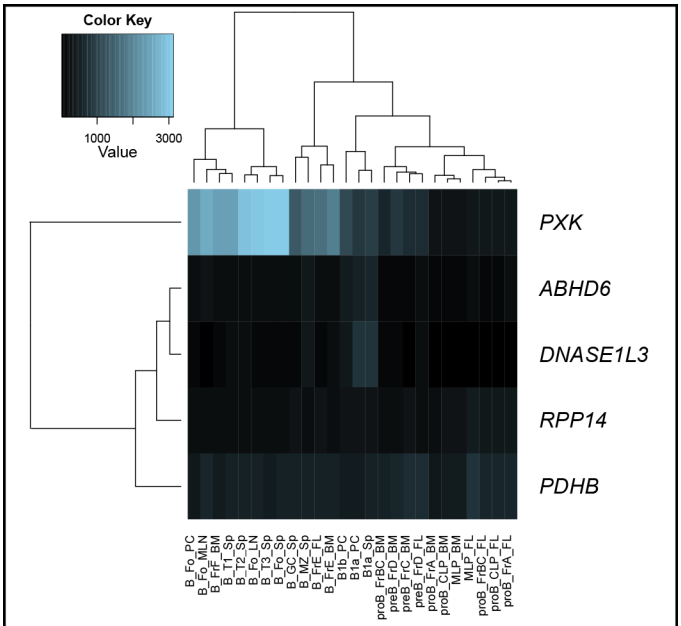


Figure S1. Expression of genes at the PXK locus in B cells. Heatmap of expression of PXK and neighboring genes in murine B cell subsets. Expression values were obtained from the ImmGen project. Heatmap generated in R using the package “heatmap.2”. The presented dataset from mice was chosen because of the availability of information from B cell subsets. Unfortunately, human data sets do not yet have this amount of information available.

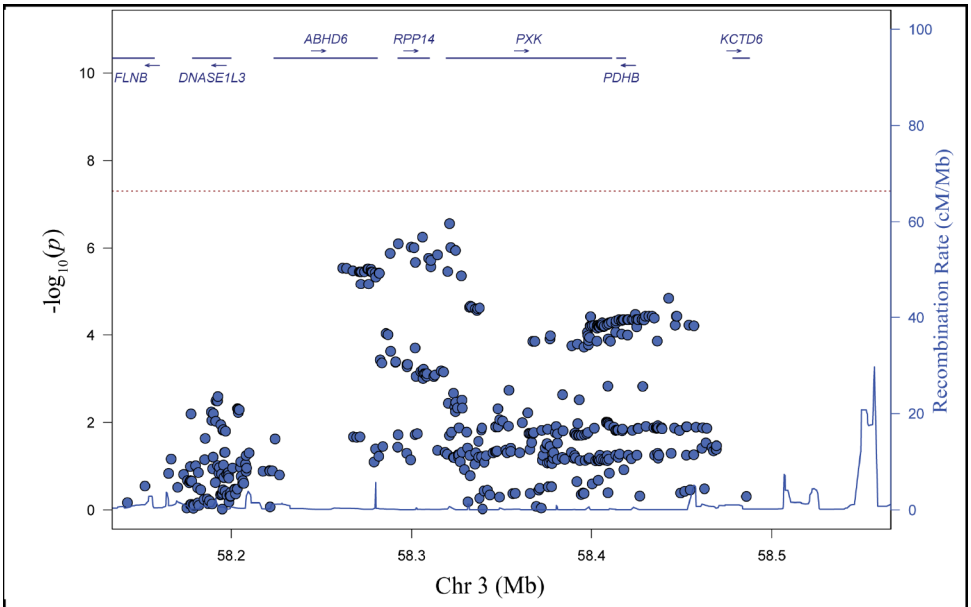


Figure S2. Step-wise logistic regression analysis of previously identified lupus-SNP in PXK. Plot of P values ($-\log_{10}$) following adjustment for genotype at SNP rs6445975. This SNP was identified in the 2008 GWAS study that identified the association of variants at the PXK locus.

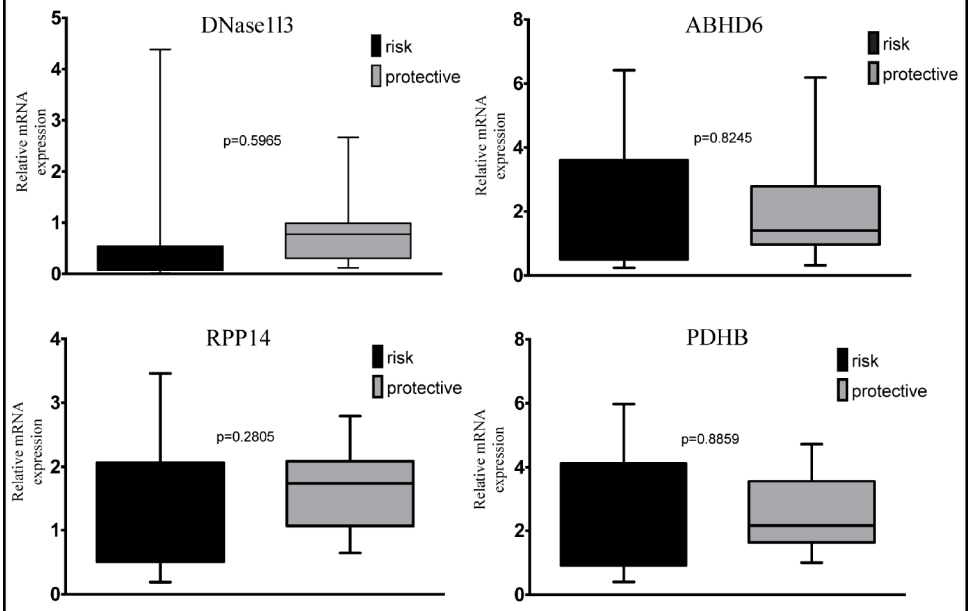


Figure S3. Expression analysis of genes at the PXK locus. Individual plots of ΔΔCt of remaining genes in the lupus-associated region of the PXK locus. 12 cell lines carrying the risk and 12 carrying the protective genotypes were assayed. Each cell line was assessed in triplicate in three independent experiments.

Chapter 4: Discussion and Conclusions

Outline of Discussion and Conclusions

results including:

Autoimmunity is a failure of homeostasis. In lupus, the clinical evidence of that loss of balance is readily apparent, but the specific pathophysiologic mechanisms leading to disruption of homeostasis are still being elucidated. Especially with the increase in data available in this era of genomic exploration, an increasing number of molecular pathways and mechanisms are being identified. The work presented in this dissertation focuses on the *PXK* locus and the genetic and biologic evidence we find supporting this locus as a potentially important player in the maintenance of balance in the immune system, and in B cells in particular.

The results of our fine-mapping study clearly support the association of a ~200kb region extending from the promoter of *PXK* through to the 3' UTR in European ancestral populations. Stepwise logistic regression supported the conclusion that there is only one independent association signal in the region, and the Bayesian analysis confirmed our results. The 95% credible set contains 172 variants in a 202 kb region. The association that we identify is confirmed by another, smaller, recent study that also demonstrated a very comparable pattern of association with similar boundaries and LD amongst the most associated variants [149]. Ideally, a fine-mapping study will allow a narrowing down of potential causal variants and highlight a specific genomic region likely to contain the casual variant. While we were able to identify a specific, well-defined region of the genome associated with SLE, which replicates the findings from another, independent, dataset, the 95% credible set still contains nearly 200 variants and the genomic region is ~ 200 kb in length. Further refinement of the genetic signal in this

locus will be crucial to fully understanding the exact genetic mechanism underlying the SLE association.

The Search for a Causal Variant

The first step we took in trying to identify the causal variant in the *PXK* region was conduct a multi-ancestral fine-mapping study. Even with the large number of subjects in our study, it is becoming apparent that even larger numbers are necessary for multi-ancestral analysis [383, 384]. Our power analysis suggests that to detect an effect of a similar size to that in the EA ancestry, we would need at least twice as many subjects. Another potential limitation of our current study was the number of variants typed in the *PXK* locus. 57 markers were included in the analysis, covering approximately 300kb of the target locus. We used imputation to increase our coverage of the region, including an additional 10,000 markers covering over 1MB of DNA in the region. Most of these markers were non-polymorphic in some samples and were eliminated from further analyses, leaving us with 835 variants total in the EA samples across the 300kb primary region. While this represents a more than 10-fold increase from our directly genotyped markers, and also represents the most complete set we have to date of the catalogued common variants in this region, we may still be missing the causal variant. One possible explanation is that the actual causal variant is rare, and only occurs in a subset of patients and yet the association in that subset is strong enough to drive the association in the entire data set. This is very unlikely given the small effect size of the association we have identified at the *PXK* locus. It could also be

that the causal variant does exist within our pool of variants, but given the limited nature of current variant annotation, we may not fully appreciate the variant.

Defining a causal variant

What is a causal variant? A causal variant is the specific genetic change that is responsible for the biological change that increases disease risk at an associated locus. Given that most disease-associated genetic variants only account for a small proportion of the absolute risk of acquiring a given complex genetic disease, any single causal variant is most likely not responsible for the entirety of the disease phenotype. Instead, in combination with other, similar genetic changes, or given specific environmental exposures, the causal variant at the associated locus increases the likelihood of acquiring the disease. In order to be casual in the context of a specific disease then, the suspected variant will lead to some biological alteration that favors disease. For autoimmune diseases such as lupus, the identified and suspected variants often lead to disease-permissive alterations to the immune system (e.g. small amplifications of the Type I IFN inflammatory signaling pathway).

In the context of the *PXK* locus, we approached the problem from two directions simultaneously- from the genetic side by attempting to refine the lupus-associated genetic signal to the smallest number of potential candidates, and the from phenotypic side, identifying a variant-linked intermediate phenotype with a potential impact on disease pathology.

Criteria for causality

A recent study outlines criteria that have been successfully used to tie discovered variants in single patients with immunodeficiency to Mendalian disorders [385]. First, the variant must be completely penetrant, meaning it occurs in individuals displaying the phenotype, and not in those that do not display the phenotype. Secondly, mechanistic studies must demonstrate a direct alteration of protein expression or function because of the genetic variant. Lastly, a causal relationship must be established between the genetic variant and the clinical phenotype either through a relevant cellular or animal model. The first requirement is obviously not applicable to complex genetic diseases as by definition, there is not a single genetic variation responsible for the clinical phenotype. In fact, even rare variants with potentially large effects may have variable penetrance in the complex disease model depending on the background genetic liability [386].

For various reasons this is likely to be more successful in Mendalian disorders where a single variant in one gene is responsible for the disease phenotype. In polygenetic disorders with environmental contribution, such as lupus, it becomes much harder to precisely implicate a functional variant with the disease. In fact, in many instances, multiple functional variants will be discovered in a given locus, especially in a region that has been extensively studied, and yet these variants may still be incidental to the specific disease risk. This appears to be true at the IRF5 locus, which has been extensively studied and multiple functional variants described and linked to autoimmune diseases such as lupus [387-389], RA [212, 390], systemic sclerosis [391], and MS

[392, 393]. We find however that both frequentist and Bayesian analysis on a larger trans-ancestral dataset reveal that the most likely causal variants are unrelated to the previously described functional variants [394]. This is also the case in the *PXK* locus, where previous studies have identified variants in DNAse1l3 that are correlated with DNase1l3 expression. In our dataset we find that these variants are not likely to be responsible for the association signal that we see. We make this conclusion based on multiple analyses, but especially on the results of the Bayesian analysis, which effectively excludes these variants from further consideration.

Investigating Rare Variants

Rare variants have been proposed to be the source for the missing heritability of genetic association studies [395] and as likely causal variants in high LD with tagging SNPs identified directly by the study. Theoretically, rare variants would be infrequently found in the population, but their disease association would be much more significant than the tagging SNPs in high LD as they would be responsible for the actual causal mutation. While this certainly may be true, initial studies utilizing exon sequencing have not demonstrated to date strong association signals amongst rare variants [396]. The exact nature of the contribution of rare variants is currently unclear and there is some argument that the amount of missing heritability is overstated, suggesting that rare variants with significant disease associations are in fact not common, and instead genetic interaction networks may be more prominent in explaining genetic inheritance of disease associated genes [248]. A further difficulty with rare variant analysis is the large

number of samples needed for adequate power given the small effect size [384, 397]. For example, one estimate for adequate sample size for rare variant analysis in Crohn's disease was 25,000, and to begin to tease out gene-gene, gene-environment interactions, that increased to as many as 500,000 subjects [384]. For complex traits, such as risk of RA (and SLE), both rare variants and accumulation of multiple common variants with small effect sizes, both in the non-coding region, appear to be important contributors to overall genetic risk of disease [397].

Differences in expression

In addition to rare variants, it may be that we have identified the causal variant, but do not yet appreciate its mechanism of action. One obvious disease-modifying phenotype to be associated with a variant (thus conferring "causality") is a change in mRNA or protein expression. We did not find any evidence in our current study, with some limitations. The first is sample size. Given the relatively small odds ratio (OR) with this association, it is likely that any change in expression will not be large. A quick power calculation suggests that to identify a 20% difference in expression with an alpha of 0.05 with 80% power, we would need 100 samples in each group, for 200 samples total. Thus, we were underpowered to detect a modest change in expression. Another limitation is that we relied on lymphoblastoid cell lines (LCLs) for our expression analysis. LCLs have been shown to differ in expression profiles from peripheral B cells [398]. The other advantage to using fresh PBMCs is the ability to measure expression in multiple cellular subsets. While we did not find any differences in PXK levels in

peripheral B cell subsets in a small sample of healthy controls, the data from the Immgen project using murine cells demonstrated the highest level of PXK in follicular B cells. It would be valuable to look at differences in PXK expression in B cell subsets from genotyped patients and controls.

A recent paper looking at SLE-associated variants in *CSK* demonstrated that the associated variant affects B cell signaling and activation, leading to changes in B cell subset distributions [224]. They found that CSK levels were highest in naïve and transitional B cells, and that the risk haplotype correlated with increased CSK expression. In naïve B cells, the risk haplotype was associated with increased B cell activation. Additionally, they also found higher levels of transitional B cells in cord blood from subjects carrying the risk haplotype. This study illustrates some of the additional information that can be gained by utilizing fresh patient samples.

Data Integration in the –omics era

Future work at identifying causal variants will increasingly depend on collecting more complete biological samples from large numbers of patients and controls in order to properly study the myriad potential consequences of genetic variation including changes in overall protein expression, protein function, expression kinetics, protein localization (both cellular and sub-cellular), alternative splice products, etc., all in the relevant biological environment. One potential approach is described below.

Deep sequencing in combination with SNP genotyping

Firstly, comprehensive genotyping information is crucial to ensure proper coverage of the risk locus. A combined approach utilizing well-designed and well-tested genotyping arrays capable of typing millions of variants per sample can be combined with deep sequencing of the identified locus and the surrounding genetic neighborhood to ensure identification of all potential variants in the region. Given the greater expense of deep sequencing when compared to array-based genotyping, sequencing can be performed on a smaller subset of individuals and the resulting data can be used as the reference panel for imputation. This allows a more cost effective approach to identifying all variants in the specific study population. An additional approach is to capitalize on GWAS identification of disease-associated loci and limit deep sequencing to those regions. This approach has been used to expand the pool of known associated variants in multiple diseases, including IBD [399], RA [397] and SLE [394, 400]. Another approach is the combination of low-coverage sequencing data with a traditional GWAS approach results in an increased effective sample size and a corresponding increase in statistical power [401].

However, even the most comprehensive genetic study will still be limited in the possible conclusions by the nature of the data available for each test subject [402]. This has been appreciated and more recent studies combine comprehensive clinical information for each subject to look beyond overall disease risk in an effort to identify genetic associations with specific phenotypes or disease subtypes [403, 404]. Functional annotation of the implicated genetic variants remains one of the key steps in assigning causality to a variant. A recent review suggested a workflow for functional

annotation including, amongst others, high-throughput expression analysis using various expression vectors, functional assays such as detecting changes to protein folding, trafficking within the cell, alteration of protein-protein interactions, or finally, more specialized functional assays specific for the protein of interest [405].

Epigenetic data and the ENCODE project: overview and applications

The Encyclopedia of DNA Elements (ENCODE) project is an effort to catalogue all functional elements within the human genome [406]. Amongst the functional elements being captured are RNA translated regions, TF-binding sites, chromatin structure and DNA methylation sites. By identifying functional elements of the genome, not only do we gain a greater understanding of what is encoded in our DNA, but it also permits the annotation of identified DNA variants with potential to impact these encoded epigenetic functional elements. Some of the potential advancements and inherent challenges of combing ENCODE data with GWAS data are reviewed in [407]. Of note, one of the greatest challenges remains the successful integration of the data. This includes development of new analytical tools and will necessarily include collaborations between teams from diverse disciplines.

One application of the ENCODE database is to reduce the list of potential causal variants at a locus by identifying the plausibly functional allele. Schaub et al. [408] describe an approach to identifying possible functional candidates. Candidate variants in a locus are aligned with the epigenetic data obtained from ENCODE. Variants are scored based on the number of overlapping functional annotation tracks, including

DNase hypersensitivity, ChIP-seq and e-QTL data. Variants that fall in regions enriched for markers of epigenetic regulation that are in high LD with the lead SNP are thus flagged as potential candidates for causal variants.

Epigenetic marks have been used successfully to characterize cancer cell lines [409]. By cataloguing the variation in both methylation and DNaseI hypersensitivity across the entire genome, the researchers were able to identify a signature that was unique to cancer cell lines. Furthermore, the epigenetic elements within the signature significantly overlapped with known cancer-associated genomic variants. This study emphasizes the power inherent in the growing body of epigenetic data to identify regions of epigenetic variability and their overlap with disease-associated genomic variants.

Leveraging RNA-seq to more fully understand sequence variation

The next level of information that would help refine the genetic associations in a disease specific manner would be expression data. RNA-seq is a powerful method allowing analysis of both overall expression and specific variations in exon usage. By combining RNA-seq data with comprehensive genotyping data, genotypic variants associated with specific expression patterns can be identified. This approach can be especially powerful when RNA-seq is performed on multiple cell types from each individual subject.

For example, researchers recently performed an RNA-seq analysis of cord blood containing hematopoietic precursors identified a novel transcription start site and novel

exon for NFIB, which they called NFIB-S. The novel, shortened form of NFIB was restricted to hematopoietic lineages, unlike the longer, canonical isoform, NFIB-L, which is broadly expressed across many tissues. By using both NFIB-S and NFIB-L in knockdown and overexpression experiments, they were able to show that NFIB-S plays a specific role in the development of megakaryocytes [410]. This study underscores the power of using RNA-seq to identify in a more thorough and complete manner differential gene expression, including alternative splice events, in specific cellular populations. Applying a similar analysis to RNA-seq data obtained from sorted lymphocyte subsets from SLE-patient and control populations, for whom genotypes are known, would allow a much more complete analysis of specific expression changes associated with SLE genotypes.

Alternative splicing of mRNA is a specific example of expression modulation at the mRNA level. Splice variation for individual genes has been shown to occur for nearly all genes and often occurs between different tissue types [411]. RNA-seq has been successfully used to quantify alternative splicing [412] and this technique has been leveraged to investigate the functional relationship between genetic risk and disease manifestations. In particular, alternative splicing has been shown to play a role in the development of autism spectrum disorders [413]. Alternative splicing has also been identified in monocytes from SLE patients when compared to healthy controls [414].

Two recent studies highlight the application of RNA-seq to lupus genetics and provide some guidance as to how a similar study might be done with *PXK*. Stone et al. examined alternative splicing of IRF5; identifying both a unique transcript signature in

subjects with SLE, which is associated with an IRF5 lupus-risk haplotype, in addition to demonstrating a difference in the biochemical stability of the various IRF5 isoforms [415]. One of the advantages of RNA-seq is the unbiased approach to variant identification. By applying next-generation sequencing to the cDNA as opposed to the traditional approach utilizing molecular cloning, a much greater number of clones can be isolated, greatly increasing the ability to detect transcribed isoforms. By using this unbiased approach, they were able to identify not only unique transcripts in patients with lupus, but were also able to identify a unique signature of 4 transcripts that were found most frequently in patients carrying the previously reported [388] IRF5 H2 risk haplotype.

Combining genotype information with expression data such as RNA-seq would provide a powerful platform for identifying potential causal variants, but would still fall short of fully validating the impact of any given single variant on the RNA-seq-defined phenotype. The first step in validation would be replication of the results using a different approach, such as real-time PCR, to replicate the findings. Cell lines could be identified that carry the risk and non-risk variant and then real-time PCR could be performed to validate a change in overall RNA expression for instance, and the subsequent impact on protein expression and function. While this is an important step in the validation process, more powerful strategies could further demonstrate the impact of an individual variant. One such strategy would be to directly mutate the identified base in a single cell line to both the risk and non-risk genotype. The strength of this approach

would be the normalization of any other background genetic or epigenetic modifiers on the phenotype.

CRISPRs as Tools for Validation and Therapy

Single base mutation in a cell line can has been achieved in multiple ways, but the CRISPR/Cas9 system has been shown to provide many advantages including the ability to be used in many immune cells given their utilization of viral transduction for introduction into the cell, and the ability to customize the insert to introduce a single base change, which is an important validation of the impact of potential causal SNP [416]. Another advantage of CRIPRs is that they can be used to generate the genotype of interest even with non-exomic mutations [417]. CRISPRs have been used to knockout genes [418], enhancer regions identified in GWAS [419], and have the potential to be used to introduce point mutations for verification of genetic association studies. They have also been used to successfully generate mice with multiple genetic mutations in one step [420].

In addition to serving as a tool to validate causal mutations, CRISPRs also have great promise as gene therapy. CRISPRs have been used to successfully repair the dystrophin gene in *mdx* mice, a model of Duchene's Muscular Dystrophy [421], and have been used to correct point mutations in cells derived from patients with sickle cell disease [422]. They have also been used to target multiple genes simultaneously in primates [423].

Another consideration is the impact of the phenotype identified by RNA-seq on disease pathobiology. The identified expression change and the specific cell type in which it occurred would obviously guide these experiments. An example of how this is being done is exemplified by a recent study that analyzed the state of the whole genome methylation status during B cell development [424]. They utilized 6 different human B cell populations from both peripheral blood and bone marrow samples, assessing the methylation state of the entire genome in each B cell population. They were able to show that methylation status, which is an indicator of areas of active transcription, occurs most frequently in naïve or precursor cell types, and that the lowest levels of methylation were found in terminally differentiated cell types, such as plasmacytes. Of note, there were unique patterns of methylation of CpGs between the various subsets, highlighting the importance of epigenetic regulation to the developmental stage of the cell. This study underscores the potential complexity in studying the impact of a candidate variant in a biological context. By looking at the impact of a given variant on epigenetic modifiers in only one population, it could be easy to miss a change that might be unique to another cellular population in an alternate developmental pathway.

Utilization of animal models in understanding the *PXK* locus

A further step in validating a potential causal variant would be to study the impact of that specific variant in an animal model of the disease. By using the aforementioned techniques, animals carrying both the risk and non-risk genotype in an otherwise

identical genetic/epigenetic environment could be generated, allowing the study of the impact of the specific variant on cellular function in vivo. The strength and importance of such an approach would be the ability to analyze the effect of a variant in the context of an intact system. In addition, perturbations to the system that would lead to a phenotype in vivo that might possibly be undetected in an in vitro setting (such as cell-cell interactions), could now be assessed.

We attempted to generate PXK knockout mice using ES cells carrying loxP sites allowing both germline knockout and cell type specific conditional knockout mice. Utilizing embryonic stem (ES) cells purchased from the International Mouse Phenotyping Consortium, part of the European Conditional Mouse Mutagenesis Program [425], we arranged for the injection of the ES cells into a carrier mouse resulting in the production of one chimeric mouse carrying the targeting construct. From our single chimeric male we generated nearly 200 pups, none of which carried the construct as a germline mutation. Unfortunately, the chimeric founder died before producing offspring carrying the germline construct. Given that we did not know the exact causal variant in the *PXK* locus, but we hypothesized that *PXK* was the effected gene, we planned to generate both germline knockout mice and a parallel mouse line capable of conditional knockout of PXK in specific cell types.

As we found a clear link between PXK and BCR internalization following PXK knockdown with shRNA, generating an animal model would provide valuable information. Animal experiments would focus first on replicating this result in vivo. We would cross our PXK conditional knockout mice carrying loxP sites flanking exons 6 and

7 of PXK with CD19-cre mice [426], which would result in cre expression only in CD19+ B cells, leading to deletion of the floxed exons of PXK and functional PXK deletion in B cells. B cells would be isolated and BCR internalization would be measured as described previously. We anticipate that tissue-specific PXK deficiency would lead to a similar decrease in BCR internalization as we saw in human immortalized B cell lines.

Further experiments would focus on defining the functional consequences of PXK depletion within the context of animal models of SLE. The initial animal model we would utilize is the pristane-induced model of SLE. While there are numerous spontaneous mouse models of lupus, the advantage of using the pristane model would be ease of induction. We could utilize the existing mice that had been generated for the previous described experiments without the need to backcross into one of the strains with spontaneous lupus.

Intraperitoneal pristane injections result in autoantibody production similar to that found in human lupus [427, 428]. In addition, these animals develop several clinical features similar to human lupus including lupus nephritis and lupus arthritis [429]. Our experiments would focus on quantifying disease following pristane-mediated induction in CD19-cre/PXK-flox mice. We would quantify disease by looking at autoantibody production including ANA and dsDNA production and by quantifying glomerulonephritis [430]. We anticipate that PXK-knockdown would result in protection from disease as the disease-associated genotype in our human study is protective.

Leveraging the Data with Complex Traits

Complex traits are by their very nature difficult to analyze and assess. A complex trait is a trait that is dependent on interactions between multiple molecular factors, including environmental and behavioral modification. Mendalian traits, in contrast, are those that can be explained by variation at a single genetic allele and thus follow Mendalian patterns of inheritance. Given the multitude of potential interactions underlying the characteristics of a single complex trait, it is apparent that working toward an understanding of these interactions and their influence on the manifestation of that trait will require immense coordination of large data sets and integration of the data in novel ways. In the current –omics era, with the plethora of large datasets emerging from the quantification of an expanding number of cellular characteristics, we are continuing to increase our capability to better understand the consequences of these interactions.

One of the next challenges will be to move beyond studies focused on single data types to those that utilize all available data, leveraging as much of the available data as possible to build more complex and complete models of the actual relationship between inherited genetic variation and the phenotypic manifestation of complex traits. The importance of looking at multiple types of data instead of focusing on one single data type is illustrated in the above review with a hypothetical example using a breast cancer model. A similar example could be made for lupus.

As we described in our earlier review [128], many of the genes so far associated with lupus participate in B cell signaling pathways. As we are able to leverage this information and integrate it with the emerging data from epigenetics [431], including DNA methylation [432, 433], expression data such as eQTLs [377], comparison to other

autoimmune diseases [434], and more, patterns will emerge that can be used to drive not just discovery, but treatment as well. A recent example of this approach exemplifying the potential inherit in combining multiple types of data is given by Okada et al. [435] for RA. In this study, they successfully combine, amongst other data, a large GWAS of greater than 100,000 individuals, eQTL analysis, pathway-based analysis, epigenetic data available for the loci, overlap between identified genes and genes with previously described immune functions, and genes with known targeted drugs, to prioritize and validate existing and newly identified associated genetic variants. In addition, they also identify newly associated genes with known target drugs already in use in other diseases as potential novel drugs that can be trialed in patients with RA.

Part of the current difficulty is that there are changes that we may not even be able to detect by observing only part of the available data. Patterns may emerge through the integration of multiple datasets that are not visible in the single data sets alone

Further Exploration of BCR signaling

One of the main modifiers of BCR signaling is the duration of time the BCR-antigen complex spends on the B cell surface. As BCR internalization has been shown to attenuate signaling, molecules that alter the rate of BCR internalization are also likely to alter the magnitude or duration of BCR-antigen induced signaling.

As both PXK knockdown and SLE-associated variants in PXK are associated with the rate of BCR internalization following BCR crosslinking, we hypothesize that

BCR signaling will be decreased in a similar manner. In order to test this hypothesis, we would first measure the impact of the SLE-associated risk variants in the *PXK* locus on constituents of the BCR signaling pathway; namely Ca^{2+} signaling and second messenger phosphorylation. Secondly, we would measure the potential of PXK to impact antigen presentation. In addition, changes in BCR internalization may affect BCR trafficking within the cell.

Calcium signaling following BCR crosslinking

For the first experiments focused on immediate signaling consequences, cell lines containing the most highly associated risk variant or not would be treated with anti-BCR as performed previously in the BCR-internalization assays. Cells would then be assayed via flow cytometry and western blot analysis for phosphorylation of second messenger molecules, including Lyn, Syk, which are early in the signaling cascade. Ca^{2+} mobilization experiments would be run in a similar but slightly different manner. The same cell lines would be used, but cells would be loaded with Indo-1 and the 405/480 nm emission ratio monitored via flow cytometry. The emission ratio of Indo-1 shifts from about 475 nm without Ca^{2+} to around 400 nm in the presence of Ca^{2+}. Thus an increase in the 405/480 nm ratio is a reliable indicator of Ca^{2+} levels [436]. Cell would be loaded with Indo-1 for 30 min at 37 deg, then a baseline ratio measured on a flow cytometry. Cells would then be crosslinked as described above and the 405/480 nm ratio monitored for 60 min to measure Ca^{2+} flux. As our hypothesis is that the delay in BCR internalization leads to a decrease in BCR-dependent signaling, we would

anticipate that both Ca^{2+} flux and second messenger phosphorylation will be decreased following PXK knockdown and in those cell lines carrying the SLE-associated allele. This would suggest that PXK facilitates BCR signaling at that the SLE-association in the region may be protective by decreasing B cell activity.

Antigen presentation by stimulated B cells

BCR internalization does not only affect BCR signaling however. It has also been shown to have an impact on antigen presentation by the B cell [437, 438]. It may be that the change in BCR internalization we observe in cell lines carrying the SLE-associated variant at the PXK locus and following PXK knockdown may also lead to changes in BCR-mediated antigen presentation. In order to determine the relationship between PXK and B cell antigen presentation, we would utilize transgenic murine B cell lines in which we would manipulate PXK. The M12g4Rd B cell line has been transfected with a BCR specific for the hapten PC [371, 439]. We would manipulate PXK expression in these cells either using virally transfected shRNA to decrease expression, as we have done previously in LCLs (Chapter 3, Figure 7), or by using the same transfection system to introduce exogenous PXK, resulting in increased protein levels. Following transfection, cells would be stimulated with PC-Ova (either the SIINFEKL or ISQAVHAAHAEINEAGR peptide), and antigen presentation measured by co-culturing PC-ova stimulated cells with ova specific DO.11.10 transgenic T cells. Activation of the PC-ova specific T cells would reflect surface antigen presentation by the M12g4Rd B

cells. Alternatively, T cell activation following antigen presentation could be quantified by determining the amount of IL-2 in the supernatant.

We hypothesize that as shRNA mediated knockdown of PXK leads to a decrease in BCR internalization that we will observe a decrease in antigen presentation by the M12g4Rd B cells following PXK knockdown. This would be evidenced by either a decrease in tetramer staining or a decrease in IL-2 amounts following stimulation with PC-Ova.

Intracellular BCR trafficking following BCR crosslinking

Alternatively, the delay in BCR internalization we have observed may instead lead to a change in the subcellular destination of the BCR-antigen complex. Changes in subcellular localization of the BCR can lead to changes in the nature of the signal [437]. In order to investigate the impact of PXK on the subcellular localization of the BCR, we would measure colocalization between the BCR and markers of intracellular compartments involved in BCR internalization. We would again utilize the LCLs derived from SLE patients and controls with known risk/nonrisk genotypes at the SLE-associated *PXK* locus and shRNA-mediated PXK knockdown in LCLs. To determine changes in BCR trafficking following internalization, cells would be crosslinked as previously described and then antibody stained in order to visualize the intracellular compartments. Antibodies associated with intracellular locations include EEA1 (early endosome) and LAMP1 (late endosome/lysosome). By measuring the colocalization between the BCR and intracellular markers over time following BCR crosslinking, we

can quantify the change in colocalization and therefore subcellular localization of the BCR following PXK manipulation.

Regulation of B cell survival and activation through the BCR

The BCR-antigen complex fails to enter the late endosome in anergic B cells [438]. This disruption in BCR trafficking is hypothesized to prevent the efficient presentation of cognate antigen to T cells as the primary site of antigen processing in B cells is the late endosome [440]. If the SLE-associated variant in the PXK locus likewise disrupts delivery of antigen to the late endosome, either through delayed internalization or alteration of the BCR transit through the B cell, this could have profound effects on both B cell fate and the subsequent immune response. The delivery of the BCR to the late endosome is not only disrupted in anergic B cells, but can actually induce anergy as restoration of BCR-antigen complex to the late endosome actually reverses anergy [438]. Additionally, rapid BCR internalization has been shown to be inhibitory in B cells, leading to downregulation of B cell signaling [441]. BCR trafficking is a key modulator of antigen presentation [442], [443].

BCR internalization is also an important regulator of BCR signaling. Blocking BCR internalization changes the gene activation profile following BCR crosslinking [355]. Phosphorylated BCR remains on the cell surface to form the signalsome responsible for the afferent signaling pathway. This signaling begins with phosphorylation of ITAMs on Igα leading to recruitment of SFTKs and signal cascade induction. However, ITAM motifs on Igα and Igβ not only serve as the site of recruitment

of second messenger molecules following phosphorylation of these motifs [444-446], but are also important for receptor internalization [447]. The ITAM motif is also recognized by AP2, which helps regulate receptor internalization through clathrin-coated pits [448]. Therefore, endocytosis of BCR complexes is dependent upon ITAMs that are not phosphorylated. So BCR internalization attenuates BCR signaling not by internalization of actively signaling (phosphorylated) receptor complexes, but instead by reducing the number of available receptors on the cell surface [336].

Following internalization, BCR-antigen complexes are sorted through the early endosome to arrive in the late endosome. This sorting is thought to be facilitated by the endosomal complex required for transport (ESCRT), as occurs in other APCs. The ESCRT machinery functions primarily by recognizing ubiquitin modifications of receptors [449, 450]. Crosslinking of BCR on the cell-surface leads to ubiquitination of the receptor complex and delivery to the ESCRT pathway, mediated at least in part by the ubiquitin ligase, Cbl-b [451].

Activation of lymphocytes typically requires at least two signals. For T cells, the integration of these two signals occurs on the cell surface through binding of co-stimulatory receptors on the cell surface of APCs [452]. B cells also require two signals, but unlike in T cells, the signal integration occurs instead in the late endosome [437, 453]. Entry into the MIIC facilitates antigen processing and MHC class II loading and is regulated by BCR signaling [454].

High affinity autoreactive B cells clones are eliminated early in B cell development through clonal deletion and receptor editing [101, 103, 455-457]. However,

in normal controls, up to 20% of mature naïve B cells are autoreactive, while in patients with SLE, that number increases to as many as 50% [458]. Notably, this large pool of self-reactive B cells does not regularly contribute to the humoral immune response. Instead, these autoreactive cells largely persist in the body in an inactive state of anergy [459-461], which is thought to be the means by which most autoreactive B cells are silenced in the periphery [462]. One mechanism by which anergy is thought to inhibit cellular activation is through the exclusions of both the BCR and the co-signaling activators such as TLR9, from the MIIC [438]. This endocytic block arrests both TLR9 and the BCR just outside the late endosome. Restoration of TLR9 and the BCR to the late endosome overcomes this block.

BCR surface expression decreases in the context of antigen recognition during B cell development [463]. IgM downregulation is proportional to the amount of antigen recognition. Interestingly, these cells (with low BCR surface expression) are not eliminated but persist in the mature B cell pool. Furthermore, anti-ANA reactive B cells are overrepresented in this population, leading to the hypothesis that these cells may represent a source for pathogenic antibodies in autoimmune disease [463].

Potential Consequences of Delayed BCR Internalization

We found that the SLE associated variants at the *PXK* locus lead to a decrease in BCR internalization. Given that BCR internalization is linked with both the B cell activation state and delivery of antigen for processing, there are multiple potential consequences of these variants. Firstly, the decrease in BCR internalization could lead

to an increase in survival signals early in development, resulting in potential autoreactive cells that are a normal part of the dormant pool in the periphery now persisting with a lower threshold for activation. These cells would usually be found in a dormant state in the periphery due to BCR downregulation and dampening of BCR signaling. Now however, with a decrease in BCR internalization, BCR signaling may be increased, leading to a lower threshold for activation and an increase in activation of autoreactive clones. One way to test this hypothesis would be to look simply at peripheral B cell activation states in individuals striated according to the SLE-associated risk allele at the PXK locus. We would expect that individuals carrying the allele associated with delayed internalization would have an increased population of activated B cells, using measures such as cell-surface expression of CD27 or other markers of B cells activation.

Another possible mechanism, however, could be an alteration of delivery of antigen to the MIIC with consequences to both cellular activation and/or antigen presentation. As stated earlier, delivery of the BCR is an important step in the regulation of anergy. Presumably, a decrease in BCR internalization would lead to a delay in BCR delivery to the MIIC. This could lead to both a disruption of antigen presentation and also to a reduction in overall activation, pushing the cell toward anergy.

Summary and Conclusion

During our analysis of the *PXK* locus, we streamlined our data analysis of SNP genotyping data sets. We developed a systematic approach, integrating rigorous

statistical testing of genetic association, complete evaluation of the target locus given the available data, imputation to fully exploit the publically available data, followed by targeted methods aimed at identifying underlying genetic structure at the locus. We used this approach to successfully confirm disease association at multiple loci, including *IRF5* [394], *PXK* [464], and *ETS1* [465]. As we adopted a unified approach to each data set, utilizing similar analytical approaches for each, we then could develop streamline our analysis across all data sets. The analysis method is described in the 3rd chapter of this dissertation. By using the same analytic approach for each locus, we could join the analysis tools using simple computer programs to facilitate processing from start to end. We were also able to analyze multiple potential genetic models at once by processing the different scenarios simultaneously, again facilitated by simple computer programs. By streamlining the initial analysis, we are able to spend less time on identifying regions of association, and can instead focus on verification and following up potential biological findings.

At the *PXK* locus, we confirmed the SLE association and identified the region of the locus most likely to contain the variant responsible for the association. In addition, we were able to show that the SLE association at the locus was also associated with a delay in BCR internalization from the cell surface. This is the first time that the function of the B cell receptor is associated with SLE-risk variants. We also showed that *PXK* levels in the cell correlate with BCR internalization. This novel finding is especially intriguing given the limited data supporting a role for PXK in receptor trafficking within the cell [141].

BCR internalization is necessary for the transportation of antigen to the MIIC to be processed and loaded onto MHCII molecules for recruitment of T cell help. BCR internalization also reduces the strength of the signalsome by reducing the number of unphosphorylated BCR receptors on the cell surface, thus dampening the overall afferent signal strength. In addition, the MIIC is the location of the integration of the co-stimulatory activation signal and thus reduction of BCR delivery to the MIIC promotes anergy.

Given these described consequences of BCR internalization, changes in BCR internalization may have multiple consequences. Delay in BCR internalization may result in delay of MIIC antigen delivery, and consequently attenuation of recruitment of T cell help. Delay in arrival of BCR to the MIIC may also result in enhancement of the propensity toward anergy development. Alternatively, the increase in surface BCR expression due to delayed internalization may lead to enhancement of signalsome activation and B cell activation secondary to BCR complex ITAM phosphorylation.

It is therefore difficult to predict the exact consequences that PXK-mediated modulation of BCR internalization may have on B cell signaling within the context of lupus pathogenesis. However, BCR internalization is central to B cell development, survival, and modulating the B cell repertoire in the periphery. Genetic variants that lead to alterations in BCR internalization, regardless of the exact mechanism, have the potential to have a profound effect on the development of autoimmune disease.

It is the combination of all the possible mechanisms- common variants, rare variants, gene-gene interactions, gene-environment interactions, and epigenetics,

coming together to form a combined genetic contribution to disease. The exact result of any given mutation in a specific subject will depend on their background genetic liability and the strength of that specific mutation within the context of that liability. A mutation may put one person over the disease threshold, but may just move another individual closer to that threshold while remaining healthy [386]. Our understanding of lupus and in particular, the SLE-association at the *PXK* locus, will be greatly enhanced in the coming years through the push for data integration. This will likely lead to not just a more complete picture of disease pathology, but also to potential targets of intervention. Treatments targeting the activation potential of B cells by modulating BCR internalization are likely to have a profound effect on B cell mediated autoimmune diseases.

Bibliography

1. Burman KD, Baker JR, Jr. (1985) Immune mechanisms in Graves' disease. Endocr Rev 6: 183-232.

2. Murray EL, Kedar S, Vedanarayanan VV (2010) Transmission of maternal muscle-specific tyrosine kinase (MuSK) to offspring: report of two cases. J Clin Neuromuscul Dis 12: 76-79.

3. Hornberger LK, Al Rajaa N (2010) Spectrum of cardiac involvement in neonatal lupus. Scand J Immunol 72: 189-197.

4. Reichlin M, Brucato A, Frank MB, Maddison PJ, McCubbin VR, Wolfson-Reichlin M, Lee LA (1994) Concentration of autoantibodies to native 60-kd Ro/SS-A and denatured 52-kd Ro/SS-A in eluates from the heart of a child who died with congenital complete heart block. Arthritis Rheum 37: 1698-1703.

5. McNeil HP, Simpson RJ, Chesterman CN, Krilis SA (1990) Anti-phospholipid antibodies are directed against a complex antigen that includes a lipid-binding inhibitor of coagulation: beta 2-glycoprotein I (apolipoprotein H). Proc Natl Acad Sci U S A 87: 4120-4124.

6. Oosting JD, Derksen RH, Bobbink IW, Hackeng TM, Bouma BN, de Groot PG (1993) Antiphospholipid antibodies directed against a combination of phospholipids with prothrombin, protein C, or protein S: an explanation for their pathogenic mechanism? Blood 81: 2618-2625.

7. Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M (1995) Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains

(CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. J Immunol 155: 1151-1164.

8. Rasooly L, Burek CL, Rose NR (1996) Iodine-induced autoimmune thyroiditis in NOD-H-2h4 mice. Clin Immunol Immunopathol 81: 287-292.

9. Constantinescu CS, Farooqi N, O'Brien K, Gran B (2011) Experimental autoimmune encephalomyelitis (EAE) as a model for multiple sclerosis (MS). Br J Pharmacol 164: 1079-1106.

10. Dubois EL, Horowitz RE, Demopoulos HB, Teplitz R (1966) NZB/NZW mice as a model of systemic lupus erythematosus. Jama 195: 285-289.

11. Haas KM, Watanabe R, Matsushita T, Nakashima H, Ishiura N, Okochi H, Fujimoto M, Tedder TF (2010) Protective and pathogenic roles for B cells during systemic autoimmunity in NZB/W F1 mice. J Immunol 184: 4789-4800.

12. Cohen PL, Eisenberg RA (1991) Lpr and gld: single gene models of systemic autoimmunity and lymphoproliferative disease. Annual Review of Immunology 9: 243-269.

13. Andrews BS, Eisenberg RA, Theofilopoulos AN, Izui S, Wilson CB, McConahey PJ, Murphy ED, Roths JB, Dixon FJ (1978) Spontaneous murine lupus-like syndromes. Clinical and immunopathological manifestations in several strains. J Exp Med 148: 1198-1215.

14. Chaparro RJ, Dilorenzo TP (2010) An update on the use of NOD mice to study autoimmune (Type 1) diabetes. Expert Rev Clin Immunol 6: 939-955.

15. Whitacre CC (2001) Sex differences in autoimmune disease. Nat Immunol 2: 777-780.

16. Frisoni L, Mcphie L, Colonna L, Sriram U, Monestier M, Gallucci S, Caricchio R (2005) Nuclear Autoantigen Translocation and Autoantibody Opsonization Lead to Increased Dendritic Cell Phagocytosis and Presentation of Nuclear Antigens: A Novel Pathogenic Pathway for Autoimmunity? The Journal of Immunology 175: 2692-2701.

17. Kuwana M, Okazaki Y, Ikeda Y (2009) Splenic macrophages maintain the anti-platelet autoimmune response via uptake of opsonized platelets in patients with immune thrombocytopenic purpura. J Thromb Haemost 7: 322-329.

18. Dempsey PW, Allison ME, Akkaraju S, Goodnow CC, Fearon DT (1996) C3d of complement as a molecular adjuvant: bridging innate and acquired immunity. Science 271: 348-350.

19. Sheerin NS, Springall T, Carroll MC, Hartley B, Sacks SH (1997) Protection against anti-glomerular basement membrane (GBM)-mediated nephritis in C3- and C4-deficient mice. Clin Exp Immunol 110: 403-409.

20. Hudson BG, Tryggvason K, Sundaramoorthy M, Neilson EG (2003) Alport's syndrome, Goodpasture's syndrome, and type IV collagen. N Engl J Med 348: 2543-2556.

21. Adams DD, Fastier FN, Howie JB, Kennedy TH, Kilpatrick JA, Stewart RD (1974) Stimulation of the human thyroid by infusions of plasma containing LATS protector. J Clin Endocrinol Metab 39: 826-832.

22. Drachman DB, Adams RN, Josifek LF, Self SG (1982) Functional activities of autoantibodies to acetylcholine receptors and the clinical severity of myasthenia gravis. N Engl J Med 307: 769-775.

23. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, Schaller JG, Talal N, Winchester RJ (1982) The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis and Rheumatism 25: 1271-1277.

24. Hochberg MC (1997) Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis and Rheumatism 40: 1725.

25. Grossman JM (2009) Lupus arthritis. Best Pract Res Clin Rheumatol 23: 495-506.

26. Labowitz R, Schumacher HR, Jr. (1971) Articular manifestations of systemic lupus erythematosus. Ann Intern Med 74: 911-921.

27. Watanabe T, Tsuchida T (1995) Classification of lupus erythematosus based upon cutaneous manifestations. Dermatological, systemic and laboratory findings in 191 patients. Dermatology 190: 277-283.

28. Nico MM, Vilela MA, Rivitti EA, Lourenco SV (2008) Oral lesions in lupus erythematosus: correlation with cutaneous lesions. Eur J Dermatol 18: 376-381.

29. Ortega LM, Schultz DR, Lenz O, Pardo V, Contreras GN (2010) Review: Lupus nephritis: pathologic features, epidemiology and a guide to therapeutic decisions. Lupus 19: 557-574.

30. Austin HA (1998) Clinical evaluation and monitoring of lupus kidney disease. Lupus 7: 618-621.

31. Danila MI, Pons-Estel GJ, Zhang J, Vila LM, Reveille JD, Alarcon GS (2009) Renal damage is the most important predictor of mortality within the damage index: data from LUMINA LXIV, a multiethnic US cohort. Rheumatology (Oxford) 48: 542-545.

32. Lapadula G, Muolo P, Semeraro F, Covelli M, Brindicci D, Cuccorese G, Francavilla A, Pipitone V (1994) Esophageal motility disorders in the rheumatic diseases: a review of 150 patients. Clin Exp Rheumatol 12: 515-521.

33. (1999) The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. Arthritis Rheum 42: 599-608.

34. Toubi E, Khamashta MA, Panarra A, Hughes GR (1995) Association of antiphospholipid antibodies with central nervous system disease in systemic lupus erythematosus. Am J Med 99: 397-401.

35. Hay EM (1994) Psychiatric disorder and cognitive impairment in SLE. Lupus 3: 145-148.

36. Hanly JG, Fisk JD, Sherwood G, Eastwood B (1994) Clinical course of cognitive dysfunction in systemic lupus erythematosus. J Rheumatol 21: 1825-1831.

37. Kozora E, Ellison MC, West S (2006) Depression, fatigue, and pain in systemic lupus erythematosus (SLE): relationship to the American College of Rheumatology SLE neuropsychological battery. Arthritis Rheum 55: 628-635.

38. Ward MM, Studenski S (1991) The time course of acute psychiatric episodes in systemic lupus erythematosus. J Rheumatol 18: 535-539.

39. Bertsias GK, Ioannidis JP, Aringer M, Bollen E, Bombardieri S, Bruce IN, Cervera R, Dalakas M, Doria A, Hanly JG, Huizinga TW, Isenberg D, Kallenberg C, Piette JC, Schneider M, Scolding N, Smolen J, Stara A, Tassiulas I, Tektonidou M, Tincani A, van Buchem MA, van Vollenhoven R, Ward M, Gordon C, Boumpas DT (2010) EULAR recommendations for the management of systemic lupus erythematosus with neuropsychiatric manifestations: report of a task force of the EULAR standing committee for clinical affairs. Ann Rheum Dis 69: 2074-2082.

40. Preble JM, Silpa-Archa S, Foster CS (2015) Ocular involvement in systemic lupus erythematosus. Curr Opin Ophthalmol 26: 540-545.

41. Newman K, Owlia MB, El-Hemaidi I, Akhtari M (2013) Management of immune cytopenias in patients with systemic lupus erythematosus - Old and new. Autoimmun Rev 12: 784-791.

42. Matsuyama T, Kuwana M, Matsumoto M, Isonishi A, Inokuma S, Fujimura Y (2009) Heterogeneous pathogenic processes of thrombotic microangiopathies in patients with connective tissue diseases. Thromb Haemost 102: 371-378.

43. Shapira Y, Weinberger A, Wysenbeek AJ (1996) Lymphadenopathy in systemic lupus erythematosus. Prevalence and relation to disease manifestations. Clin Rheumatol 15: 335-338.

44. Barile-Fabris L, Hernandez-Cabrera MF, Barragan-Garfias JA (2014) Vasculitis in systemic lupus erythematosus. Curr Rheumatol Rep 16: 440.

45. Satoh M, Chan EK, Ho LA, Rose KM, Parks CG, Cohn RD, Jusko TA, Walker NJ, Germolec DR, Whitt IZ, Crockett PW, Pauley BA, Chan JY, Ross SJ, Birnbaum

LS, Zeldin DC, Miller FW (2012) Prevalence and sociodemographic correlates of antinuclear antibodies in the United States. Arthritis Rheum 64: 2319-2327.

46. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA, Harley JB (2003) Development of autoantibodies before the clinical onset of systemic lupus erythematosus. New England Journal of Medicine 349: 1526-1533.

47. Rose NR (2008) Predictors of autoimmune disease: autoantibodies and beyond. Autoimmunity 41: 419-428.

48. Mariz HA, Sato EI, Barbosa SH, Rodrigues SH, Dellavance A, Andrade LE (2011) Pattern on the antinuclear antibody-HEp-2 test is a critical parameter for discriminating antinuclear antibody-positive healthy individuals and patients with autoimmune rheumatic diseases. Arthritis Rheum 63: 191-200.

49. Meroni PL, Schur PH (2010) ANA screening: an old test with new recommendations. Ann Rheum Dis 69: 1420-1422.

50. Slater CA, Davis RB, Shmerling RH (1996) Antinuclear antibody testing. A study of clinical utility. Arch Intern Med 156: 1421-1425.

51. Tan EM, Chan EK, Sullivan KF, Rubin RL (1988) Antinuclear antibodies (ANAs): diagnostically specific immune markers and clues toward the understanding of systemic autoimmunity. Clin Immunol Immunopathol 47: 121-141.

52. Tan EM, Feltkamp TE, Smolen JS, Butcher B, Dawkins R, Fritzler MJ, Gordon T, Hardin JA, Kalden JR, Lahita RG, Maini RN, McDougal JS, Rothfield NF,

Smeenk RJ, Takasaki Y, Wiik A, Wilson MR, Koziol JA (1997) Range of antinuclear antibodies in "healthy" individuals. Arthritis Rheum 40: 1601-1611.

53. Benito-Garcia E, Schur PH, Lahita R (2004) Guidelines for immunologic laboratory testing in the rheumatic diseases: anti-Sm and anti-RNP antibody tests. Arthritis Rheum 51: 1030-1044.

54. Rahman A, Isenberg DA (2008) Systemic lupus erythematosus. N Engl J Med 358: 929-939.

55. ter Borg EJ, Horst G, Hummel EJ, Limburg PC, Kallenberg CG (1990) Measurement of increases in anti-double-stranded DNA antibody levels as a predictor of disease exacerbation in systemic lupus erythematosus. A long-term, prospective study. Arthritis Rheum 33: 634-643.

56. Bootsma H, Spronk P, Derksen R, de Boer G, Wolters-Dicke H, Hermans J, Limburg P, Gmelig-Meyling F, Kater L, Kallenberg C (1995) Prevention of relapses in systemic lupus erythematosus. Lancet 345: 1595-1599.

57. Tseng CE, Buyon JP, Kim M, Belmont HM, Mackay M, Diamond B, Marder G, Rosenthal P, Haines K, Ilie V, Abramson SB (2006) The effect of moderate-dose corticosteroids in preventing severe flares in patients with serologically active, but clinically stable, systemic lupus erythematosus: findings of a prospective, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 54: 3623-3632.

58. Amoura Z, Koutouzov S, Chabre H, Cacoub P, Amoura I, Musset L, Bach JF, Piette JC (2000) Presence of antinucleosome autoantibodies in a restricted set of

connective tissue diseases: antinucleosome antibodies of the IgG3 subclass are markers of renal pathogenicity in systemic lupus erythematosus. Arthritis Rheum 43: 76-84.

59. Buyon JP, Clancy RM (2003) Maternal autoantibodies and congenital heart block: mediators, markers, and therapeutic approach. Semin Arthritis Rheum 33: 140-154.

60. Sontheimer RD, Maddison PJ, Reichlin M, Jordon RE, Stastny P, Gilliam JN (1982) Serologic and HLA associations in subacute cutaneous lupus erythematosus, a clinical subset of lupus erythematosus. Ann Intern Med 97: 664-671.

61. McCarty GA, Harley JB, Reichlin M (1993) A distinctive autoantibody profile in black female patients with lupus nephritis. Arthritis Rheum 36: 1560-1565.

62. Yoshio T, Onda K, Nara H, Minota S (2006) Association of IgG anti-NR2 glutamate receptor antibodies in cerebrospinal fluid with neuropsychiatric systemic lupus erythematosus. Arthritis Rheum 54: 675-678.

63. Lapteva L, Nowak M, Yarboro CH, Takada K, Roebuck-Spencer T, Weickert T, Bleiberg J, Rosenstein D, Pao M, Patronas N, Steele S, Manzano M, van der Veen JW, Lipsky PE, Marenco S, Wesley R, Volpe B, Diamond B, Illei GG (2006) Anti-N-methyl-D-aspartate receptor antibodies, cognitive dysfunction, and depression in systemic lupus erythematosus. Arthritis Rheum 54: 2505-2514.

64. Alarcon-Segovia D, Deleze M, Oria CV, Sanchez-Guerrero J, Gomez-Pacheco L, Cabiedes J, Fernandez L, Ponce de Leon S (1989) Antiphospholipid antibodies and the antiphospholipid syndrome in systemic lupus erythematosus. A

prospective analysis of 500 consecutive patients. Medicine (Baltimore) 68: 353-365.

65. Mason LJ, Ravirajan CT, Rahman A, Putterman C, Isenberg DA (2004) Is alpha-actinin a target for pathogenic anti-DNA antibodies in lupus nephritis? Arthritis Rheum 50: 866-870.

66. Becker-Merok A, Kalaaji M, Haugbro K, Nikolaisen C, Nilsen K, Rekvig OP, Nossent JC (2006) Alpha-actinin-binding antibodies in relation to systemic lupus erythematosus and lupus nephritis. Arthritis Res Ther 8: R162.

67. Siegert CE, Daha MR, Swaak AJ, van der Voort EA, Breedveld FC (1993) The relationship between serum titers of autoantibodies to C1q and age in the general population and in patients with systemic lupus erythematosus. Clin Immunol Immunopathol 67: 204-209.

68. Wardemann H (2003) Predominant Autoantibody Production by Early Human B Cell Precursors. Science 301: 1374-1377.

69. Ignatowicz L, Kappler J, Marrack P (1996) The repertoire of T cells shaped by a single MHC/peptide ligand. Cell 84: 521-529.

70. Zerrahn J, Held W, Raulet DH (1997) The MHC reactivity of the T cell repertoire prior to positive and negative selection. Cell 88: 627-636.

71. Hayter SM, Cook MC (2012) Updated assessment of the prevalence, spectrum and case definition of autoimmune disease. Autoimmun Rev 11: 754-765.

72. Goodnow CC, Sprent J, de St Groth BF, Vinuesa CG (2005) Cellular and genetic mechanisms of self tolerance and autoimmunity. Nature Cell Biology 435: 590-597.

73. Takahashi T, Tanaka M, Brannan CI, Jenkins NA, Copeland NG, Suda T, Nagata S (1994) Generalized lymphoproliferative disease in mice, caused by a point mutation in the Fas ligand. Cell 76: 969-976.

74. Watanabe-Fukunaga R, Brannan CI, Copeland NG, Jenkins NA, Nagata S (1992) Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. Nature 356: 314-317.

75. Rieux-Laucat F, Le Deist F, Hivroz C, Roberts IA, Debatin KM, Fischer A, de Villartay JP (1995) Mutations in Fas associated with human lymphoproliferative syndrome and autoimmunity. Science 268: 1347-1349.

76. Izui S, Kelley VE, Masuda K, Yoshida H, Roths JB, Murphy ED (1984) Induction of various autoantibodies by mutant gene lpr in several strains of mice. Journal of immunology (Baltimore, Md : 1950) 133: 227-233.

77. Watson ML, Rao JK, Gilkeson GS, Ruiz P (1992) Genetic analysis of MRL-lpr mice: relationship of the Fas apoptosis gene to disease manifestations and renal disease-modifying loci. The Journal of ….

78. Nagata S, Golstein P (1995) The Fas death factor. Science.

79. Strasser A, Bouillet P (2003) The control of apoptosis in lymphocyte selection. Immunol Rev 193: 82-92.

80. Nagamine K, Peterson P, Scott HS, Kudoh J, Minoshima S, Heino M, Krohn KJ, Lalioti MD, Mullis PE, Antonarakis SE, Kawasaki K, Asakawa S, Ito F, Shimizu N (1997) Positional cloning of the APECED gene. Nat Genet 17: 393-398.

81. Aaltonen J, Björses P, Perheentupa J, Finnish-German AC (1997) An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. Nature 17: 399-403.

82. Liston A, Lesage S, Wilson J, Peltonen L, Goodnow CC (2003) Aire regulates negative selection of organ-specific T cells. Nature Immunology 4: 350-354.

83. Gavanescu I, Benoist C (2008) B cells are required for Aire-deficient mice to develop multi-organ autoinflammation: a therapeutic approach for APECED patients. Proceedings of the ….

84. Popler J, Alimohammadi M, Kämpe O, Dalin F, Dishop MK, Barker JM, Moriarty-Kelsey M, Soep JB, Deterding RR (2011) Autoimmune polyendocrine syndrome type 1: Utility of KCNRG autoantibodies as a marker of active pulmonary disease and successful treatment with rituximab. Pediatric Pulmonology 47: 84-87.

85. Navarra SV, Guzman RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, Li EK, Thomas M, Kim HY, Leon MG, Tanasescu C, Nasonov E, Lan JL, Pineda L, Zhong ZJ, Freimuth W, Petri MA (2011) Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 377: 721-731.

86. Stohl W, Hiepe F, Latinis KM, Thomas M, Scheinberg MA, Clarke A, Aranow C, Wellborne FR, Abud-Mendoza C, Hough DR, Pineda L, Migone TS, Zhong ZJ,

Freimuth WW, Chatham WW (2012) Belimumab reduces autoantibodies, normalizes low complement levels, and reduces select B cell populations in patients with systemic lupus erythematosus. Arthritis Rheum 64: 2328-2337.

87. Dooley MA, Houssiau F, Aranow C, D'Cruz DP, Askanase A, Roth DA, Zhong ZJ, Cooper S, Freimuth WW, Ginzler EM (2013) Effect of belimumab treatment on renal outcomes: results from the phase 3 belimumab clinical trials in patients with SLE. Lupus 22: 63-72.

88. Anolik J, Sanz I, Looney RJ (2003) B cell depletion therapy in systemic lupus erythematosus. Curr Rheumatol Rep 5: 350-356.

89. Eisenberg R, Albert D (2006) B-cell targeted therapies in rheumatoid arthritis and systemic lupus erythematosus. Nat Clin Pract Rheumatol 2: 20-27.

90. Gunnarsson I, Sundelin B, Jonsdottir T, Jacobson SH, Henriksson EW, van Vollenhoven RF (2007) Histopathologic and clinical outcome of rituximab treatment in patients with cyclophosphamide-resistant proliferative lupus nephritis. Arthritis Rheum 56: 1263-1272.

91. Looney RJ, Anolik JH, Campbell D, Felgar RE, Young F, Arend LJ, Sloand JA, Rosenblatt J, Sanz Ia (2004) B cell depletion as a novel treatment for systemic lupus erythematosus: A phase I/II dose-escalation trial of rituximab. Arthritis and Rheumatism 50: 2580-2589.

92. Patel DD (2002) B cell-ablative therapy for the treatment of autoimmune diseases. Arthritis Rheum 46: 1984-1985.

93. Silverman GJ, Weisman S (2003) Rituximab therapy and autoimmune disorders: prospects for anti-B cell therapy. Arthritis Rheum 48: 1484-1492.

94. Tanaka Y, Yamamoto K, Takeuchi T, Nishimoto N, Miyasaka N, Sumida T, Shima Y, Takada K, Matsumoto I, Saito K, Koike T (2007) A multicenter phase I/II trial of rituximab for refractory systemic lupus erythematosus. Mod Rheumatol 17: 191-197.

95. Rovin BH, Furie R, Latinis K, Looney RJ, Fervenza FC, Sanchez-Guerrero J, Maciuca R, Zhang D, Garg JP, Brunetta P, Appel G, Group ftLI (2012) Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: The lupus nephritis assessment with rituximab (LUNAR) study. Arthritis and Rheumatism: n/a-n/a.

96. Beckwith H, Lightstone L (2014) Rituximab in Systemic Lupus Erythematosus and Lupus Nephritis. Nephron Clinical Practice 128: 250-254.

97. Weidenbusch M, Rommele C, Schrottle A, Anders HJ (2013) Beyond the LUNAR trial. Efficacy of rituximab in refractory lupus nephritis. Nephrology Dialysis Transplantation 28: 106-111.

98. Grammer AC, Slota R, Fischer R, Gur H, Girschick H, Yarboro C, Illei GG, Lipsky PE (2003) Abnormal germinal center reactions in systemic lupus erythematosus demonstrated by blockade of CD154-CD40 interactions. Journal of Clinical Investigation 112: 1506-1520.

99. Fuxa M, Skok JA (2007) Transcriptional regulation in early B cell development. Curr Opin Immunol 19: 129-136.

100. LeBien TW, Tedder TF (2008) B lymphocytes: how they develop and function. Blood 112: 1570-1580.

101. Gay D, Saunders T, Camper S, Weigert M (1993) Receptor editing: an approach by autoreactive B cells to escape tolerance. The Journal of experimental medicine 177: 999-1008.

102. Radic MZ, Erikson J, Litwin S, Weigert M (1993) B lymphocytes may escape tolerance by revising their antigen receptors. J Exp Med 177: 1165-1173.

103. Tiegs SL, Russell DM, Nemazee D (1993) Receptor editing in self-reactive bone marrow B cells. The Journal of experimental medicine 177: 1009-1020.

104. Pelanda R, Schwers S, Sonoda E, Torres RM, Nemazee D, Rajewsky K (1997) Receptor editing in a transgenic mouse model: site, efficiency, and role in B cell tolerance and antibody diversification. Immunity 7: 765-775.

105. Huang H, Kearney JF, Grusby MJ, Benoist C, Mathis D (2006) Induction of tolerance in arthritogenic B cells with receptors of differing affinity for self-antigen. Proc Natl Acad Sci U S A 103: 3734-3739.

106. Karasuyama H, Kudo A, Melchers F (1990) The proteins encoded by the VpreB and lambda 5 pre-B cell-specific genes can associate with each other and with mu heavy chain. J Exp Med 172: 969-972.

107. Tsubata T, Reth M (1990) The products of pre-B cell-specific genes (lambda 5 and VpreB) and the immunoglobulin mu chain form a complex that is transported onto the cell surface. J Exp Med 172: 973-976.

108. Chen C, Nagy Z, Prak EL, Weigert M (1995) Immunoglobulin heavy chain gene replacement: a mechanism of receptor editing. Immunity 3: 747-755.

109. Sims GP, Ettinger R, Shirota Y, Yarboro CH, Illei GG, Lipsky PE (2005) Identification and characterization of circulating human transitional B cells. Blood 105: 4390-4398.

110. Benitez A, Weldon AJ, Tatosyan L, Velkuru V, Lee S, Milford TA, Francis OL, Hsu S, Nazeri K, Casiano CM, Schneider R, Gonzalez J, Su RJ, Baez I, Colburn K, Moldovan I, Payne KJ (2014) Differences in mouse and human nonmemory B cell pools. J Immunol 192: 4610-4619.

111. Petro JB (2002) Transitional Type 1 and 2 B Lymphocyte Subsets Are Differentially Responsive to Antigen Receptor Signaling. Journal of Biological Chemistry 277: 48009-48019.

112. Su TT, Rawlings DJ (2002) Transitional B Lymphocyte Subsets Operate as Distinct Checkpoints in Murine Splenic B Cell Development. The Journal of Immunology 168: 2101-2110.

113. Sater RA, Sandel PC, Monroe JG (1998) B cell receptor-induced apoptosis in primary transitional murine B cells: signaling requirements and modulation by T cell help. International Immunology 10: 1673-1682.

114. Yurasov S, Tiller T, Tsuiji M, Velinzon K, Pascual V, Wardemann H, Nussenzweig MC (2006) Persistent expression of autoantibodies in SLE patients in remission. Journal of Experimental Medicine 203: 2255-2261.

115. Chang N-H, McKenzie T, Bonventi G, Landolt-Marticorena C, Fortin PR, Gladman D, Urowitz M, Wither JE (2008) Expanded population of activated antigen-engaged cells within the naive B cell compartment of patients with systemic lupus erythematosus. Journal of immunology (Baltimore, Md : 1950) 180: 1276-1284.

116. Cuss AK, Avery DT, Cannons JL, Yu LJ, Nichols KE, Shaw PJ, Tangye SG (2006) Expansion of functionally immature transitional B cells is associated with human-immunodeficient states characterized by impaired humoral immunity. Journal of immunology (Baltimore, Md : 1950) 176: 1506-1516.

117. Vossenkamper A, Lutalo PM, Spencer J (2012) Translational Mini-Review Series on B cell subsets in disease. Transitional B cells in systemic lupus erythematosus and Sjogren's syndrome: clinical implications and effects of B cell-targeted therapies. Clin Exp Immunol 167: 7-14.

118. Costenbader KH, Kim DJ, Peerzada J, Lockman S, Nobles-Knight D, Petri M, Karlson EW (2004) Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis. Arthritis Rheum 50: 849-857.

119. Rodriguez Huerta MD, Trujillo-Martin MM, Rua-Figueroa I, Cuellar-Pompa L, Quiros-Lopez R, Serrano-Aguilar P (2015) Healthy lifestyle habits for patients with systemic lupus erythematosus: A systemic review. Semin Arthritis Rheum.

120. Kiyohara C, Washio M, Horiuchi T, Asami T, Ide S, Atsumi T, Kobashi G, Takahashi H, Tada Y (2012) Risk modification by CYP1A1 and GSTM1 polymorphisms in the association of cigarette smoking and systemic lupus erythematosus in a Japanese population. Scand J Rheumatol 41: 103-109.

121. Lehmann P, Holzle E, Kind P, Goerz G, Plewig G (1990) Experimental reproduction of skin lesions in lupus erythematosus by UVA and UVB radiation. J Am Acad Dermatol 22: 181-187.

122. Nyberg F, Fransson J, Stephansson E (2000) Proliferation and effects of UVA irradiation in cultured fibroblasts from lesions in cutaneous lupus erythematosus. Exp Dermatol 9: 53-57.

123. Avalos-Diaz E, Alvarado-Flores E, Herrera-Esparza R (1999) UV-A irradiation induces transcription of IL-6 and TNF alpha genes in human keratinocytes and dermal fibroblasts. Rev Rhum Engl Ed 66: 13-19.

124. Casciola-Rosen LA, Anhalt G, Rosen A (1994) Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes. J Exp Med 179: 1317-1330.

125. Finckh A, Cooper GS, Chibnik LB, Costenbader KH, Watts J, Pankey H, Fraser PA, Karlson EW (2006) Occupational silica and solvent exposures and risk of systemic lupus erythematosus in urban women. Arthritis Rheum 54: 3648-3654.

126. Lu-Fritts PY, Kottyan LC, James JA, Xie C, Buckholz JM, Pinney SM, Harley JB (2014) Systemic lupus erythematosus is associated with uranium exposure in a community living near a uranium processing plant: A nested case-control study. Arthritis Rheumatol.

127. Parks CG, De Roos AJ (2014) Pesticides, chemical and industrial exposures in relation to systemic lupus erythematosus. Lupus 23: 527-536.

128. Vaughn SE, Kottyan LC, Munroe ME, Harley JB (2012) Genetic susceptibility to lupus: the biological basis of genetic risk found in B cell signaling pathways. Journal of Leukocyte Biology 92: 577-591.

129. Halder I, Shriver M, Thomas M, Fernandez JR, Frudakis T (2008) A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. Hum Mutat 29: 648-658.

130. Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM (2004) Design and analysis of admixture mapping studies. Am J Hum Genet 74: 965-978.

131. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics 38: 904-909.

132. Zou X, Qiu G, Chen C, Wu M, Hu Y, Zheng H, Li X, Gu S, Ji C, Mao Y (2005) Expression pattern and subcellular localization of five splice isoforms of human PXK. International journal of molecular medicine 16: 701-707.

133. Ponting CP (1996) Novel domains in NADPH oxidase subunits, sorting nexins, and PtdIns 3-kinases: binding partners of SH3 domains? Protein science : a publication of the Protein Society 5: 2353-2357.

134. Xu Y, Seet LF, Hanson B, HONG W (2001) The Phox homology (PX) domain, a new player in phosphoinositide signalling. Biochemical Journal 360: 513-530.

135. Teasdale RD, Collins BM (2011) Insights into the PX (phox-homology) domain and SNX (sorting nexin) protein families: structures, functions and roles in disease. Biochemical Journal 441: 39-59.

136. Vollert CS (2004) The Phox Homology (PX) Domain Protein Interaction Network in Yeast. Molecular & Cellular Proteomics 3: 1053-1064.

137. van Weering JR, Verkade P, Cullen PJ (2010) SNX-BAR proteins in phosphoinositide-mediated, tubular-based endosomal sorting. Semin Cell Dev Biol 21: 371-380.

138. van Kerkhof P, Lee J, McCormick L, Tetrault E, Lu W, Schoenfish M, Oorschot V, Strous GJ, Klumperman J, Bu G (2005) Sorting nexin 17 facilitates LRP recycling in the early endosome. Embo j 24: 2851-2861.

139. Joubert L, Hanson B, Barthet G, Sebben M, Claeysen S, Hong W, Marin P, Dumuis A, Bockaert J (2004) New sorting nexin (SNX27) and NHERF specifically interact with the 5-HT4a receptor splice variant: roles in receptor targeting. J Cell Sci 117: 5367-5379.

140. Choi JH, Hong WP, Kim MJ, Kim JH, Ryu SH, Suh PG (2004) Sorting nexin 16 regulates EGF receptor trafficking by phosphatidylinositol-3-phosphate interaction with the Phox domain. J Cell Sci 117: 4209-4218.

141. Takeuchi H, Takeuchi T, Gao J, Cantley LC, Hirata M (2010) Characterization of PXK as a Protein Involved in Epidermal Growth Factor Receptor Trafficking. Molecular and Cellular Biology 30: 1689-1702.

142. Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD (2008) Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nature Genetics 40: 204-210.

143. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X, Ortmann W, Kosoy R, Ferreira RC, Nordmark G, Gunnarsson I, Svenungsson E, Padyukov L, Sturfelt G, Jönsen A, Bengtsson AA, Rantapää-Dahlqvist S, Baechler EC, Brown EE, Alarcón GS, Edberg JC, Ramsey-Goldman R, Mcgwin G, Reveille JD, Vilá LM, Kimberly RP, Manzi S, Petri MA, Lee A, Gregersen PK, Seldin MF, Rönnblom L, Criswell LA, Syvänen A-C, Behrens TW, Graham RR (2009) A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nature Genetics 41: 1228-1233.

144. Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Sebastiani G, Ruzickova S, Jose Santos M, Papasteriades C, Marchini M, Skopouli FN, Suarez A, Blanco FJ, D'Alfonso S, Bijl M, Carreira P, Witte T, Migliaresi S, Gomez-Reino JJ, Gonzalez A, Collections tECoSD (2009) Replication of recently identified systemic lupus erythematosus genetic associations: a case–control study. Arthritis Research & Therapy 11: R69.

145. Martin JE, Assassi S, Diaz-Gallo L-M, Broen JC, Simeon CP, Castellví I, Vicente-Rabaneda E, Fonollosa V, Ortego-Centeno N, Gonzalez-Gay MA, Espinosa G, Carreira P, Group SS, consortium S, group USSG, BIOLUPUS, Camps M, Sabio

JM, D'Alfonso S, Vonk MC, Voskuyl AE, Schuerwegh AJ, Kreuter A, Witte T, Riemekasten G, Hunzelmann N, Airó P, Beretta L, Scorza R, Lunardi C, Van Laar J, Chee MM, Worthington J, Herrick A, Denton C, Fonseca C, Tan FK, Arnett F, Zhou X, Reveille JD, Gorlova O, Koeleman BPC, Radstake TRDJ, Vyse T, Mayes MD, Alarcón-Riquelme ME, Martin J (2013) A systemic sclerosis and systemic lupus erythematosus pan-meta-GWAS reveals new shared susceptibility loci. Human Molecular Genetics 22: 4021-4029.

146. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, Li Y, Kurreeman FAS, Zhernakova A, Hinks A, Guiducci C, Chen R, Alfredsson L, Amos CI, Ardlie KG, Barton A, Bowes J, Brouwer E, Burtt NP, Catanese JJ, Coblyn J, Coenen MJH, Costenbader KH, Criswell LA, Crusius JBA, Cui J, de Bakker PIW, De Jager PL, Ding B, Emery P, Flynn E, Harrison P, Hocking LJ, Huizinga TWJ, Kastner DL, Ke X, Lee AT, Liu X, Martin P, Morgan AW, Padyukov L, Posthumus MD, Radstake TRDJ, Reid DM, Seielstad M, Seldin MF, Shadick NA, Steer S, Tak PP, Thomson W, van der Helm-van Mil AHM, van der Horst-Bruinsma IE, van der Schoot CE, van Riel PLCM, Weinblatt ME, Wilson AG, Wolbink GJ, Wordsworth BP, Wijmenga C, Karlson EW, Toes REM, de Vries N, Begovich AB, Worthington J, Siminovitch KA, Gregersen PK, Klareskog L, Plenge RM (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nature Genetics 42: 508-514.

147. Yu B, Wu Q, Chen Y, Li P, Shao Y, Zhang J, Zhong Q, Peng X, Yang H, Hu X, Chen B, Guan M, Zhang W, Wan J (2011) Polymorphisms of PXK are associated

with autoantibody production, but not disease risk, of systemic lupus erythematosus in Chinese mainland population. Lupus 20: 23-27.

148. Kim E-M, Bang S-Y, Kim IL, Shin HD, Park BL, Lee H-S, Bae S-C (2011) Different genetic effect of PXK on systemic lupus erythematosus in the Korean population. Rheumatology International 32: 277-280.

149. Oparina NY, Delgado-Vega AM, Martinez-Bueno M, Magro-Checa C, Fernandez C, Castro RO, Pons-Estel BA, D'Alfonso S, Sebastiani GD, Witte T, Lauwerys BR, Endreffy E, Kovacs L, Escudero A, Lopez-Pedrera C, Vasconcelos C, da Silva BM, Frostegard J, Truedsson L, Martin J, Raya E, Ortego-Centeno N, de Los Angeles Aguirre M, de Ramon Garrido E, Palma MJ, Alarcon-Riquelme ME, Kozyrev SV (2015) PXK locus in systemic lupus erythematosus: fine mapping and functional analysis reveals novel susceptibility gene ABHD6. Ann Rheum Dis 74: e14.

150. Alarcón-Segovia D, Alarcón-Riquelme ME, Cardiel MH, Caeiro F, Massardo L, Villa AR, Pons-Estel BA, (GLADEL) GLdEdLE (2005) Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis and Rheumatism 52: 1138-1147.

151. Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, Walker A, Mack TM (1992) A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis and Rheumatism 35: 311-318.

152. Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, Chang M, Ramos P, Baechler EC, Batliwalla FM, Novitzke J, Williams AH, Gillett C, Rodine P, Graham RR, Ardlie KG, Gaffney PM, Moser KL, Petri M, Begovich AB, Gregersen PK, Behrens TW (2004) Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. Am J Hum Genet 75: 504-507.

153. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD, Nath SK, Guthridge JM, Cobb BL, Mirel DB, Marion MC, Williams AH, Divers J, Wang W, Frank SG, Namjou B, Gabriel SB, Lee AT, Gregersen PK, Behrens TW, Taylor KE, Fernando M, Zidovetzki R, Gaffney PM, Edberg JC, Rioux JD, Ojwang JO, James JA, Merrill JT, Gilkeson GS, Seldin MF, Yin H, Baechler EC, Li QZ, Wakeland EK, Bruner GR, Kaufman KM, Kelly JA (2008) Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 40: 204-210.

154. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X, Ortmann W, Kosoy R, Ferreira RC, Nordmark G, Gunnarsson I, Svenungsson E, Padyukov L, Sturfelt G, Jonsen A, Bengtsson AA, Rantapaa-Dahlqvist S, Baechler EC, Brown EE, Alarcon GS, Edberg JC, Ramsey-Goldman R, McGwin G, Jr., Reveille JD, Vila LM, Kimberly RP, Manzi S, Petri MA, Lee A, Gregersen PK, Seldin MF, Ronnblom L, Criswell LA, Syvanen AC, Behrens TW, Graham RR (2009) A large-

scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet 41: 1228-1233.

155. Budarf ML, Goyette P, Boucher G, Lian J, Graham RR, Claudio JO, Hudson T, Gladman D, Clarke AE, Pope JE, Peschken C, Smith CD, Hanly J, Rich E, Boire G, Barr SG, Zummer M, Fortin PR, Wither J, Rioux JD (2011) A targeted association study in systemic lupus erythematosus identifies multiple susceptibility alleles. Genes Immun 12: 51-58.

156. Morris DL, Roberts AL, Witherden AS, Tarzi R, Barros P, Whittaker JC, Cook TH, Aitman TJ, Vyse TJ (2010) Evidence for both copy number and allelic (NA1/NA2) risk at the FCGR3B locus in systemic lupus erythematosus. Eur J Hum Genet 18: 1027-1031.

157. Karassa FB, Trikalinos TA, Ioannidis JP (2002) Role of the Fcgamma receptor IIa polymorphism in susceptibility to systemic lupus erythematosus and lupus nephritis: a meta-analysis. Arthritis Rheum 46: 1563-1571.

158. Karassa FB, Trikalinos TA, Ioannidis JP (2003) The Fc gamma RIIIA-F158 allele is a risk factor for the development of lupus nephritis: a meta-analysis. Kidney Int 63: 1475-1482.

159. McKinney C, Merriman TR (2012) Meta-analysis confirms a role for deletion in FCGR3B in autoimmune phenotypes. Hum Mol Genet 21: 2370-2376.

160. Zidan HE, Sabbah NA, Hagrass HA, Tantawy EA, El-Shahawy EE, Nageeb GS, Abdul-Sattar AB (2014) Association of FcgammaRIIB and FcgammaRIIA R131H

gene polymorphisms with renal involvement in Egyptian systemic lupus erythematosus patients. Mol Biol Rep 41: 733-739.

161. Yu B, Chen Y, Wu Q, Li P, Shao Y, Zhang J, Zhong Q, Peng X, Yang H, Hu X, Chen B, Guan M, Wan J, Zhang W (2011) The association between single-nucleotide polymorphisms of NCF2 and systemic lupus erythematosus in Chinese mainland population. Clin Rheumatol 30: 521-527.

162. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burtt NP, Guiducci C, Parkin M, Gates C, Plenge RM, Behrens TW, Wither JE, Rioux JD, Fortin PR, Graham DC, Wong AK, Vyse TJ, Daly MJ, Altshuler D, Moser KL, Gaffney PM (2008) Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet 40: 1059-1061.

163. Jacob CO, Eisenstein M, Dinauer MC, Ming W, Liu Q, John S, Quismorio FP, Jr., Reiff A, Myones BL, Kaufman KM, McCurdy D, Harley JB, Silverman E, Kimberly RP, Vyse TJ, Gaffney PM, Moser KL, Klein-Gitelman M, Wagner-Weiner L, Langefeld CD, Armstrong DL, Zidovetzki R (2012) Lupus-associated causal mutation in neutrophil cytosolic factor 2 (NCF2) brings unique insights to the structure and function of NADPH oxidase. Proc Natl Acad Sci U S A 109: E59-67.

164. Cunninghame Graham DS, Morris DL, Bhangale TR, Criswell LA, Syvanen AC, Ronnblom L, Behrens TW, Graham RR, Vyse TJ (2011) Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus. PLoS Genet 7: e1002341.

165. Kim-Howard X, Sun C, Molineros JE, Maiti AK, Chandru H, Adler A, Wiley GB, Kaufman KM, Kottyan L, Guthridge JM, Rasmussen A, Kelly J, Sanchez E, Raj P, Li QZ, Bang SY, Lee HS, Kim TH, Kang YM, Suh CH, Chung WT, Park YB, Choe JY, Shim SC, Lee SS, Han BG, Olsen NJ, Karp DR, Moser K, Pons-Estel BA, Wakeland EK, James JA, Harley JB, Bae SC, Gaffney PM, Alarcon-Riquelme M, on behalf of G, Looger LL, Nath SK (2013) Allelic heterogeneity in NCF2 associated with systemic lupus erythematosus (SLE) susceptibility across four ethnic populations. Hum Mol Genet.

166. Edberg JC, Wu J, Langefeld CD, Brown EE, Marion MC, McGwin G, Jr., Petri M, Ramsey-Goldman R, Reveille JD, Frank SG, Kaufman KM, Harley JB, Alarcon GS, Kimberly RP (2008) Genetic variation in the CRP promoter: association with systemic lupus erythematosus. Hum Mol Genet 17: 1147-1155.

167. Sanchez E, Comeau ME, Freedman BI, Kelly JA, Kaufman KM, Langefeld CD, Brown EE, Alarcon GS, Kimberly RP, Edberg JC, Ramsey-Goldman R, Petri M, Reveille JD, Vila LM, Merrill JT, Tsao BP, Kamen DL, Gilkeson GS, James JA, Vyse TJ, Gaffney PM, Jacob CO, Niewold TB, Richardson BC, Harley JB, Alarcon-Riquelme ME, Sawalha AH (2011) Identification of novel genetic susceptibility loci in African American lupus patients in a candidate gene association study. Arthritis Rheum 63: 3493-3501.

168. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, Xu JH, Cai ZM, Huang W, Zhao GP, Xie HF, Fang H, Lu QJ, Li XP, Pan YF, Deng DQ, Zeng FQ, Ye ZZ, Zhang XY, Wang QW, Hao F, Ma L, Zuo XB, Zhou FS, Du WH, Cheng YL, Yang JQ,

Shen SK, Li J, Sheng YJ, Zuo XX, Zhu WF, Gao F, Zhang PL, Guo Q, Li B, Gao M, Xiao FL, Quan C, Zhang C, Zhang Z, Zhu KJ, Li Y, Hu DY, Lu WS, Huang JL, Liu SX, Li H, Ren YQ, Wang ZX, Yang CJ, Wang PG, Zhou WM, Lv YM, Zhang AP, Zhang SQ, Lin D, Low HQ, Shen M, Zhai ZF, Wang Y, Zhang FY, Yang S, Liu JJ, Zhang XJ (2009) Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet 41: 1234-1237.

169. Ramos PS, Criswell LA, Moser KL, Comeau ME, Williams AH, Pajewski NM, Chung SA, Graham RR, Zidovetzki R, Kelly JA, Kaufman KM, Jacob CO, Vyse TJ, Tsao BP, Kimberly RP, Gaffney PM, Alarcon-Riquelme ME, Harley JB, Langefeld CD (2011) A Comprehensive Analysis of Shared Loci between Systemic Lupus Erythematosus (SLE) and Sixteen Autoimmune Diseases Reveals Limited Genetic Overlap. PLoS Genet 7: e1002406.

170. Yu HH, Liu PH, Lin YC, Chen WJ, Lee JH, Wang LC, Yang YH, Chiang BL (2010) Interleukin 4 and STAT6 gene polymorphisms are associated with systemic lupus erythematosus in Chinese patients. Lupus 19: 1219-1228.

171. Sakurai D, Zhao J, Deng Y, Kelly JA, Brown EE, Harley JB, Bae SC, Alarcomicronn-Riquelme ME, Biolupus, networks G, Edberg JC, Kimberly RP, Ramsey-Goldman R, Petri MA, Reveille JD, Vila LM, Alarcon GS, Kaufman KM, Vyse TJ, Jacob CO, Gaffney PM, Sivils KM, James JA, Kamen DL, Gilkeson GS, Niewold TB, Merrill JT, Scofield RH, Criswell LA, Stevens AM, Boackle SA, Kim JH, Choi J, Pons-Estel BA, Argentine Collaborative G, Freedman BI, Anaya JM,

Martin J, Yu CY, Chang DM, Song YW, Langefeld CD, Chen W, Grossman JM, Cantor RM, Hahn BH, Tsao BP (2013) Preferential binding to Elk-1 by SLE-associated IL10 risk allele upregulates IL10 expression. PLoS Genet 9: e1003870.

172. Schotte H, Willeke P, Becker H, Poggemeyer J, Gaubitz M, Schmidt H, Schluter B (2014) Association of extended interleukin-10 promoter haplotypes with disease susceptibility and manifestations in German patients with systemic lupus erythematosus. Lupus 23: 378-385.

173. Bowness P, Davies KA, Norsworthy PJ, Athanassiou P, Taylor-Wiedeman J, Borysiewicz LK, Meyer PA, Walport MJ (1994) Hereditary C1q deficiency and systemic lupus erythematosus. QJM 87: 455-464.

174. Nishino H, Shibuya K, Nishida Y, Mushimoto M (1981) Lupus erythematosus-like syndrome with selective complete deficiency of C1q. Ann Intern Med 95: 322-324.

175. Suzuki Y, Ogura Y, Otsubo O, Akagi K, Fujita T (1992) Selective deficiency of C1s associated with a systemic lupus erythematosus-like syndrome. Report of a case. Arthritis Rheum 35: 576-579.

176. Walport MJ, Davies KA, Morley BJ, Botto M (1997) Complement deficiency and autoimmunity. Ann N Y Acad Sci 815: 267-281.

177. Cao CW, Li P, Luan HX, Chen W, Li CH, Hu CJ, Zhang SL, Zeng XF, Zhang FC, Li YZ, Zeng CQ (2012) Association study of C1qA polymorphisms with systemic lupus erythematosus in a Han population. Lupus 21: 502-507.

178. Yang W, Tang H, Zhang Y, Tang X, Zhang J, Sun L, Yang J, Cui Y, Zhang L, Hirankarn N, Cheng H, Pan HF, Gao J, Lee TL, Sheng Y, Lau CS, Li Y, Chan TM, Yin X, Ying D, Lu Q, Leung AM, Zuo X, Chen X, Tong KL, Zhou F, Diao Q, Tse NK, Xie H, Mok CC, Hao F, Wong SN, Shi B, Lee KW, Hui Y, Ho MH, Liang B, Lee PP, Cui H, Guo Q, Chung BH, Pu X, Liu Q, Zhang X, Zhang C, Chong CY, Fang H, Wong RW, Sun Y, Mok MY, Li XP, Avihingsanon Y, Zhai Z, Rianthavorn P, Deekajorndej T, Suphapeetiporn K, Gao F, Shotelersuk V, Kang X, Ying SK, Zhang L, Wong WH, Zhu D, Fung SK, Zeng F, Lai WM, Wong CM, Ng IO, Garcia-Barcelo MM, Cherny SS, Shen N, Tam PK, Sham PC, Ye DQ, Yang S, Zhang X, Lau YL (2013) Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. Am J Hum Genet 92: 41-51.

179. Molineros JE, Maiti AK, Sun C, Looger LL, Han S, Kim-Howard X, Glenn S, Adler A, Kelly JA, Niewold TB, Gilkeson GS, Brown EE, Alarcon GS, Edberg JC, Petri M, Ramsey-Goldman R, Reveille JD, Vila LM, Freedman BI, Tsao BP, Criswell LA, Jacob CO, Moore JH, Vyse TJ, Langefeld CL, Guthridge JM, Gaffney PM, Moser KL, Scofield RH, Alarcon-Riquelme ME, Network B, Williams SM, Merrill JT, James JA, Kaufman KM, Kimberly RP, Harley JB, Nath SK (2013) Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production. PLoS Genet 9: e1003222.

180. Espeli M, Clatworthy MR, Bokers S, Lawlor KE, Cutler AJ, Kontgen F, Lyons PA, Smith KG (2012) Analysis of a wild mouse promoter variant reveals a novel role for FcgammaRIIb in the control of the germinal center and autoimmunity. J Exp Med 209: 2307-2319.

181. Abelson AK, Delgado-Vega AM, Kozyrev SV, Sanchez E, Velazquez-Cruz R, Eriksson N, Wojcik J, Linga Reddy MV, Lima G, D'Alfonso S, Migliaresi S, Baca V, Orozco L, Witte T, Ortego-Centeno N, Abderrahim H, Pons-Estel BA, Gutierrez C, Suarez A, Gonzalez-Escribano MF, Martin J, Alarcon-Riquelme ME (2009) STAT4 associates with systemic lupus erythematosus through two independent effects that correlate with gene expression and act additively with IRF5 to increase risk. Ann Rheum Dis 68: 1746-1753.

182. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, de Bakker PI, Le JM, Lee HS, Batliwalla F, Li W, Masters SL, Booty MG, Carulli JP, Padyukov L, Alfredsson L, Klareskog L, Chen WV, Amos CI, Criswell LA, Seldin MF, Kastner DL, Gregersen PK (2007) STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 357: 977-986.

183. Yuan H, Feng JB, Pan HF, Qiu LX, Li LH, Zhang N, Ye DQ (2010) A meta-analysis of the association of STAT4 polymorphism with systemic lupus erythematosus. Mod Rheumatol 20: 257-262.

184. Namjou B, Sestak AL, Armstrong DL, Zidovetzki R, Kelly JA, Jacob N, Ciobanu V, Kaufman KM, Ojwang JO, Ziegler J, Quismorio FP, Jr., Reiff A, Myones BL, Guthridge JM, Nath SK, Bruner GR, Mehrian-Shai R, Silverman E, Klein-

Gitelman M, McCurdy D, Wagner-Weiner L, Nocton JJ, Putterman C, Bae SC, Kim YJ, Petri M, Reveille JD, Vyse TJ, Gilkeson GS, Kamen DL, Alarcon-Riquelme ME, Gaffney PM, Moser KL, Merrill JT, Scofield RH, James JA, Langefeld CD, Harley JB, Jacob CO (2009) High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups. Arthritis Rheum 60: 1085-1095.

185. Suarez-Gestal M, Calaza M, Endreffy E, Pullmann R, Ordi-Ros J, Sebastiani GD, Ruzickova S, Jose Santos M, Papasteriades C, Marchini M, Skopouli FN, Suarez A, Blanco FJ, D'Alfonso S, Bijl M, Carreira P, Witte T, Migliaresi S, Gomez-Reino JJ, Gonzalez A (2009) Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. Arthritis Res Ther 11: R69.

186. Okada Y, Shimane K, Kochi Y, Tahira T, Suzuki A, Higasa K, Takahashi A, Horita T, Atsumi T, Ishii T, Okamoto A, Fujio K, Hirakata M, Amano H, Kondo Y, Ito S, Takada K, Mimori A, Saito K, Kamachi M, Kawaguchi Y, Ikari K, Mohammed OW, Matsuda K, Terao C, Ohmura K, Myouzen K, Hosono N, Tsunoda T, Nishimoto N, Mimori T, Matsuda F, Tanaka Y, Sumida T, Yamanaka H, Takasaki Y, Koike T, Horiuchi T, Hayashi K, Kubo M, Kamatani N, Yamada R, Nakamura Y, Yamamoto K (2012) A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese. PLoS Genet 8: e1002455.

187. Zhou TB, Jiang ZP, Qin YH, Zhou JF (2014) Association of STAT4 gene polymorphism with systemic lupus erythematosus / lupus nephritis risk. Nephrology (Carlton).

188. Oparina NY, Delgado-Vega AM, Martinez-Bueno M, Magro-Checa C, Fernandez C, Castro RO, Pons-Estel BA, D'Alfonso S, Sebastiani GD, Witte T, Lauwerys BR, Endreffy E, Kovacs L, Escudero A, Lopez-Pedrera C, Vasconcelos C, da Silva BM, Frostegard J, Truedsson L, Martin J, Raya E, Ortego-Centeno N, de Los Angeles Aguirre M, de Ramon Garrido E, Palma MJ, Alarcon-Riquelme ME, Kozyrev SV (2014) PXK locus in systemic lupus erythematosus: fine mapping and functional analysis reveals novel susceptibility gene ABHD6. Ann Rheum Dis.

189. Namjou B, Kothari PH, Kelly JA, Glenn SB, Ojwang JO, Adler A, Alarcon-Riquelme ME, Gallant CJ, Boackle SA, Criswell LA, Kimberly RP, Brown E, Edberg J, Stevens AM, Jacob CO, Tsao BP, Gilkeson GS, Kamen DL, Merrill JT, Petri M, Goldman RR, Vila LM, Anaya JM, Niewold TB, Martin J, Pons-Estel BA, Sabio JM, Callejas JL, Vyse TJ, Bae SC, Perrino FW, Freedman BI, Scofield RH, Moser KL, Gaffney PM, James JA, Langefeld CD, Kaufman KM, Harley JB, Atkinson JP (2011) Evaluation of the TREX1 gene in a large multi-ancestral lupus cohort. Genes Immun 12: 270-279.

190. Lee-Kirsch MA, Gong M, Chowdhury D, Senenko L, Engel K, Lee YA, de Silva U, Bailey SL, Witte T, Vyse TJ, Kere J, Pfeiffer C, Harvey S, Wong A, Koskenmies S, Hummel O, Rohde K, Schmidt RE, Dominiczak AF, Gahr M, Hollis T, Perrino

FW, Lieberman J, Hubner N (2007) Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. Nat Genet 39: 1065-1067.

191. Lessard CJ, Adrianto I, Ice JA, Wiley GB, Kelly JA, Glenn SB, Adler AJ, Li H, Rasmussen A, Williams AH, Ziegler J, Comeau ME, Marion M, Wakeland BE, Liang C, Ramos PS, Grundahl KM, Gallant CJ, Alarcon-Riquelme ME, Alarcon GS, Anaya JM, Bae SC, Boackle SA, Brown EE, Chang DM, Cho SK, Criswell LA, Edberg JC, Freedman BI, Gilkeson GS, Jacob CO, James JA, Kamen DL, Kimberly RP, Kim JH, Martin J, Merrill JT, Niewold TB, Park SY, Petri MA, Pons-Estel BA, Ramsey-Goldman R, Reveille JD, Scofield RH, Song YW, Stevens AM, Tsao BP, Vila LM, Vyse TJ, Yu CY, Guthridge JM, Kaufman KM, Harley JB, Wakeland EK, Langefeld CD, Gaffney PM, Montgomery CG, Moser KL (2012) Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. Am J Hum Genet 90: 648-660.

192. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E, Gunnarsson I, Svenungsson E, Sturfelt G, Jonsen A, Truedsson L, Pons-Estel BA, Witte T, D'Alfonso S, Barizzone N, Danieli MG, Gutierrez C, Suarez A, Junker P, Laustrup H, Gonzalez-Escribano MF, Martin J, Abderrahim H, Alarcon-Riquelme ME (2008) Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet 40: 211-216.

193. Fan Y, Tao JH, Zhang LP, Li LH, Ye DQ (2011) The association between BANK1 and TNFAIP3 gene polymorphisms and systemic lupus erythematosus: a meta-analysis. Int J Immunogenet 38: 151-159.

194. Hughes T, Kim-Howard X, Kelly JA, Kaufman KM, Langefeld CD, Ziegler J, Sanchez E, Kimberly RP, Edberg JC, Ramsey-Goldman R, Petri M, Reveille JD, Martin J, Brown EE, Vila LM, Alarcon GS, James JA, Gilkeson GS, Moser KL, Gaffney PM, Merrill JT, Vyse TJ, Alarcon-Riquelme ME, Nath SK, Harley JB, Sawalha AH (2011) Fine-mapping and transethnic genotyping establish IL2/IL21 genetic association with lupus and localize this genetic effect to IL21. Arthritis Rheum 63: 1689-1697.

195. Lan Y, Luo B, Wang JL, Jiang YW, Wei YS (2014) The association of interleukin-21 polymorphisms with interleukin-21 serum levels and risk of systemic lupus erythematosus. Gene 538: 94-98.

196. Zhong H, Li XL, Li M, Hao LX, Chen RW, Xiang K, Qi XB, Ma RZ, Su B (2011) Replicated associations of TNFAIP3, TNIP1 and ETS1 with systemic lupus erythematosus in a southwestern Chinese population. Arthritis Res Ther 13: R186.

197. Kawasaki A, Ito S, Furukawa H, Hayashi T, Goto D, Matsumoto I, Kusaoi M, Ohashi J, Graham RR, Matsuta K, Behrens TW, Tohma S, Takasaki Y, Hashimoto H, Sumida T, Tsuchiya N (2010) Association of TNFAIP3 interacting protein 1, TNIP1 with systemic lupus erythematosus in a Japanese population: a case-control association study. Arthritis Res Ther 12: R174.

198. He CF, Liu YS, Cheng YL, Gao JP, Pan TM, Han JW, Quan C, Sun LD, Zheng HF, Zuo XB, Xu SX, Sheng YJ, Yao S, Hu WL, Li Y, Yu ZY, Yin XY, Zhang XJ, Cui Y, Yang S (2010) TNIP1, SLC15A4, ETS1, RasGRP3 and IKZF1 are associated with clinical features of systemic lupus erythematosus in a Chinese Han population. Lupus 19: 1181-1186.

199. Lofgren SE, Frostegard J, Truedsson L, Pons-Estel BA, D'Alfonso S, Witte T, Lauwerys BR, Endreffy E, Kovacs L, Vasconcelos C, Martins da Silva B, Kozyrev SV, Alarcon-Riquelme ME (2012) Genetic association of miRNA-146a with systemic lupus erythematosus in Europeans through decreased expression of the gene. Genes Immun.

200. Luo X, Yang W, Ye DQ, Cui H, Zhang Y, Hirankarn N, Qian X, Tang Y, Lau YL, de Vries N, Tak PP, Tsao BP, Shen N (2011) A functional variant in microRNA-146a promoter modulates its expression and confers disease risk for systemic lupus erythematosus. PLoS Genet 7: e1002128.

201. Lofgren SE, Frostegard J, Truedsson L, Pons-Estel BA, D'Alfonso S, Witte T, Lauwerys BR, Endreffy E, Kovacs L, Vasconcelos C, Martins da Silva B, Kozyrev SV, Alarcon-Riquelme ME (2012) Genetic association of miRNA-146a with systemic lupus erythematosus in Europeans through decreased expression of the gene. Genes Immun 13: 268-274.

202. Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J, Noble JA, Taylor KE, Quach DL, Chung SA, Kelly JA, Moser KL, Behrens TW, Seldin MF, Thomson G, Harley JB, Gaffney PM, Criswell LA (2009) High-density SNP

screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS Genet 5: e1000696.

203. Morris DL, Taylor KE, Fernando MM, Nititham J, Alarcon-Riquelme ME, Barcellos LF, Behrens TW, Cotsapas C, Gaffney PM, Graham RR, Pons-Estel BA, Gregersen PK, Harley JB, Hauser SL, Hom G, Langefeld CD, Noble JA, Rioux JD, Seldin MF, Criswell LA, Vyse TJ (2012) Unraveling Multiple MHC Gene Associations with Systemic Lupus Erythematosus: Model Choice Indicates a Role for HLA Alleles and Non-HLA Genes in Europeans. Am J Hum Genet 91: 778-793.

204. Fernando MM, Freudenberg J, Lee A, Morris DL, Boteva L, Rhodes B, Gonzalez-Escribano MF, Lopez-Nevot MA, Navarra SV, Gregersen PK, Martin J, Vyse TJ (2012) Transancestral mapping of the MHC region in systemic lupus erythematosus identifies new independent and interacting loci at MSH5, HLA-DPB1 and HLA-G. Ann Rheum Dis 71: 777-784.

205. Ruiz-Narvaez EA, Fraser PA, Palmer JR, Cupples LA, Reich D, Wang YA, Rioux JD, Rosenberg L (2011) MHC region and risk of systemic lupus erythematosus in African American women. Hum Genet 130: 807-815.

206. Zhang Y, Yang W, Mok CC, Chan TM, Wong RW, Mok MY, Lee KW, Wong SN, Leung AM, Lee TL, Ho MH, Lee PP, Wong WH, Yang J, Zhang J, Wong CM, Ng IO, Garcia-Barcelo MM, Cherny SS, Tam PK, Sham PC, Lau CS, Lau YL (2011) Two missense variants in UHRF1BP1 are independently associated with

systemic lupus erythematosus in Hong Kong Chinese. Genes Immun 12: 231-234.

207. Adrianto I, Wen F, Templeton A, Wiley G, King JB, Lessard CJ, Bates JS, Hu Y, Kelly JA, Kaufman KM, Guthridge JM, Alarcon-Riquelme ME, Anaya JM, Bae SC, Bang SY, Boackle SA, Brown EE, Petri MA, Gallant C, Ramsey-Goldman R, Reveille JD, Vila LM, Criswell LA, Edberg JC, Freedman BI, Gregersen PK, Gilkeson GS, Jacob CO, James JA, Kamen DL, Kimberly RP, Martin J, Merrill JT, Niewold TB, Park SY, Pons-Estel BA, Scofield RH, Stevens AM, Tsao BP, Vyse TJ, Langefeld CD, Harley JB, Moser KL, Webb CF, Humphrey MB, Montgomery CG, Gaffney PM (2011) Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. Nat Genet 43: 253-258.

208. Lodolce JP, Kolodziej LE, Rhee L, Kariuki SN, Franek BS, McGreal NM, Logsdon MF, Bartulis SJ, Perera MA, Ellis NA, Adams EJ, Hanauer SB, Jolly M, Niewold TB, Boone DL (2010) African-derived genetic polymorphisms in TNFAIP3 mediate risk for autoimmunity. J Immunol 184: 7001-7009.

209. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Kamboh MI, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok PY, Criswell LA (2008) Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 40: 1062-1064.

210. Wang S, Wen F, Wiley GB, Kinter MT, Gaffney PM (2013) An enhancer element harboring variants associated with systemic lupus erythematosus engages the TNFAIP3 promoter to influence A20 expression. PLoS Genet 9: e1003750.

211. Sigurdsson S, Goring HH, Kristjansdottir G, Milani L, Nordmark G, Sandling JK, Eloranta ML, Feng D, Sangster-Guity N, Gunnarsson I, Svenungsson E, Sturfelt G, Jonsen A, Truedsson L, Barnes BJ, Alm G, Ronnblom L, Syvanen AC (2008) Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus. Hum Mol Genet 17: 872-881.

212. Sigurdsson S, Padyukov L, Kurreeman FA, Liljedahl U, Wiman AC, Alfredsson L, Toes R, Ronnelid J, Klareskog L, Huizinga TW, Alm G, Syvanen AC, Ronnblom L (2007) Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum 56: 2202-2210.

213. Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, Jonsen A, Rantapaa-Dahlqvist S, Moller B, Kere J, Koskenmies S, Widen E, Eloranta ML, Julkunen H, Kristjansdottir H, Steinsson K, Alm G, Ronnblom L, Syvanen AC (2005) Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet 76: 528-537.

214. Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, Ortmann WA, Koeuth T, Gonzalez Escribano MF, Pons-Estel B, Petri M, Daly M, Gregersen PK, Martin J, Altshuler D, Behrens TW, Alarcon-Riquelme ME (2006)

A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet 38: 550-555.

215. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA, Ferreira RC, Pant PV, Ballinger DG, Kosoy R, Demirci FY, Kamboh MI, Kao AH, Tian C, Gunnarsson I, Bengtsson AA, Rantapaa-Dahlqvist S, Petri M, Manzi S, Seldin MF, Ronnblom L, Syvanen AC, Criswell LA, Gregersen PK, Behrens TW (2008) Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 358: 900-909.

216. Delgado-Vega AM, Dozmorov MG, Quiros MB, Wu YY, Martinez-Garcia B, Kozyrev SV, Frostegard J, Truedsson L, de Ramon E, Gonzalez-Escribano MF, Ortego-Centeno N, Pons-Estel BA, D'Alfonso S, Sebastiani GD, Witte T, Lauwerys BR, Endreffy E, Kovacs L, Vasconcelos C, da Silva BM, Wren JD, Martin J, Castillejo-Lopez C, Alarcon-Riquelme ME (2012) Fine mapping and conditional analysis identify a new mutation in the autoimmunity susceptibility gene BLK that leads to reduced half-life of the BLK protein. Ann Rheum Dis 71: 1219-1226.

217. Chen Y, Wu Q, Shao Y, Zhang J, Guan M, Wan J, Yu B, Zhang W (2012) Identify the association between polymorphisms of BLK and systemic lupus erythematosus through unlabelled probe-based high-resolution melting analysis. Int J Immunogenet 39: 321-327.

218. Lu R, Vidal GS, Kelly JA, Delgado-Vega AM, Howard XK, Macwana SR, Dominguez N, Klein W, Burrell C, Harley IT, Kaufman KM, Bruner GR, Moser KL, Gaffney PM, Gilkeson GS, Wakeland EK, Li QZ, Langefeld CD, Marion MC, Divers J, Alarcon GS, Brown EE, Kimberly RP, Edberg JC, Ramsey-Goldman R, Reveille JD, McGwin G, Jr., Vila LM, Petri MA, Bae SC, Cho SK, Bang SY, Kim I, Choi CB, Martin J, Vyse TJ, Merrill JT, Harley JB, Alarcon-Riquelme ME, Nath SK, James JA, Guthridge JM (2009) Genetic associations of LYN with systemic lupus erythematosus. Genes Immun 10: 397-403.

219. Yang W, Shen N, Ye DQ, Liu Q, Zhang Y, Qian XX, Hirankarn N, Ying D, Pan HF, Mok CC, Chan TM, Wong RW, Lee KW, Mok MY, Wong SN, Leung AM, Li XP, Avihingsanon Y, Wong CM, Lee TL, Ho MH, Lee PP, Chang YK, Li PH, Li RJ, Zhang L, Wong WH, Ng IO, Lau CS, Sham PC, Lau YL (2010) Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. PLoS Genet 6: e1000841.

220. Zhao H, Yang W, Qiu R, Li J, Xin Q, Wang X, Feng Y, Shan S, Liu Y, Gong Y, Liu Q (2012) An intronic variant associated with systemic lupus erythematosus changes the binding affinity of Yinyang1 to downregulate WDFY4. Genes Immun 13: 536-542.

221. Lessard CJ, Adrianto I, Kelly JA, Kaufman KM, Grundahl KM, Adler A, Williams AH, Gallant CJ, Anaya JM, Bae SC, Boackle SA, Brown EE, Chang DM, Criswell LA, Edberg JC, Freedman BI, Gregersen PK, Gilkeson GS, Jacob CO, James JA, Kamen DL, Kimberly RP, Martin J, Merrill JT, Niewold TB, Park SY, Petri MA,

Pons-Estel BA, Ramsey-Goldman R, Reveille JD, Song YW, Stevens AM, Tsao BP, Vila LM, Vyse TJ, Yu CY, Guthridge JM, Bruner GR, Langefeld CD, Montgomery C, Harley JB, Scofield RH, Gaffney PM, Moser KL (2011) Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study. Am J Hum Genet 88: 83-91.

222. Zikherman J, Hermiston M, Steiner D, Hasegawa K, Chan A, Weiss A (2009) PTPN22 Deficiency Cooperates with the CD45 E613R Allele to Break Tolerance on a Non-Autoimmune Background. The Journal of Immunology 182: 4093-4106.

223. Yang J, Yang W, Hirankarn N, Ye DQ, Zhang Y, Pan HF, Mok CC, Chan TM, Wong RW, Mok MY, Lee KW, Wong SN, Leung AM, Li XP, Avihingsanon Y, Rianthavorn P, Deekajorndej T, Suphapeetiporn K, Shotelersuk V, Baum L, Kwan P, Lee TL, Ho MH, Lee PP, Wong WH, Zeng S, Zhang J, Wong CM, Ng IO, Garcia-Barcelo MM, Cherny SS, Tam PK, Sham PC, Lau CS, Lau YL (2011) ELF1 is associated with systemic lupus erythematosus in Asian populations. Hum Mol Genet 20: 601-607.

224. Manjarrez-Orduno N, Marasco E, Chung SA, Katz MS, Kiridly JF, Simpfendorfer KR, Freudenberg J, Ballard DH, Nashi E, Hopkins TJ, Cunninghame Graham DS, Lee AT, Coenen MJ, Franke B, Swinkels DW, Graham RR, Kimberly RP, Gaffney PM, Vyse TJ, Behrens TW, Criswell LA, Diamond B, Gregersen PK (2012) CSK regulatory polymorphism is associated with systemic lupus erythematosus and influences B-cell signaling and activation. Nat Genet 44: 1227-1230.

225. Molineros JE, Kim-Howard X, Deshmukh H, Jacob CO, Harley JB, Nath SK (2009) Admixture in Hispanic Americans: its impact on ITGAM association and implications for admixture mapping in SLE. Genes Immun 10: 539-545.

226. Fossati-Jimack L, Ling GS, Cortini A, Szajna M, Malik TH, McDonald JU, Pickering MC, Cook HT, Taylor PR, Botto M (2013) Phagocytosis is the main CR3-mediated function affected by the lupus-associated variant of CD11b in human myeloid cells. PLoS One 8: e57082.

227. Toller-Kawahisa JE, Vigato-Ferreira IC, Pancoto JA, Mendes-Junior CT, Martinez EZ, Palomino GM, Louzada-Junior P, Donadi EA, Del Lama JE, Marzocchi-Machado CM (2014) The variant of CD11b, rs1143679 within ITGAM, is associated with systemic lupus erythematosus and clinical manifestations in Brazilian patients. Hum Immunol 75: 119-123.

228. Sheng YJ, Gao JP, Li J, Han JW, Xu Q, Hu WL, Pan TM, Cheng YL, Yu ZY, Ni C, Yao S, He CF, Liu YS, Li Y, Ge HM, Xiao FL, Sun LD, Yang S, Zhang XJ (2011) Follow-up study identifies two novel susceptibility loci PRKCB and 8p11.21 for systemic lupus erythematosus. Rheumatology (Oxford) 50: 682-688.

229. Zhang Z, Cheng Y, Zhou X, Li Y, Gao J, Han J, Quan C, He S, Lv Y, Hu D, Zhu K, Sun L, Yang S, Zhang X (2011) Polymorphisms at 16p13 are associated with systemic lupus erythematosus in the Chinese population. J Med Genet 48: 69-72.

230. Vazgiourakis VM, Zervou MI, Choulaki C, Bertsias G, Melissourgaki M, Yilmaz N, Sidiropoulos P, Plant D, Trouw LA, Toes RE, Kardassis D, Yavuz S, Boumpas DT, Goulielmos GN (2011) A common SNP in the CD40 region is associated with

systemic lupus erythematosus and correlates with altered CD40 expression: implications for the pathogenesis. Ann Rheum Dis 70: 2184-2190.

231. Shen N, Fu Q, Deng Y, Qian X, Zhao J, Kaufman KM, Wu YL, Yu CY, Tang Y, Chen JY, Yang W, Wong M, Kawasaki A, Tsuchiya N, Sumida T, Kawaguchi Y, Howe HS, Mok MY, Bang SY, Liu FL, Chang DM, Takasaki Y, Hashimoto H, Harley JB, Guthridge JM, Grossman JM, Cantor RM, Song YW, Bae SC, Chen S, Hahn BH, Lau YL, Tsao BP (2010) Sex-specific association of X-linked Toll-like receptor 7 (TLR7) with male systemic lupus erythematosus. Proc Natl Acad Sci U S A 107: 15838-15843.

232. Lee YH, Lee HS, Choi SJ, Ji JD, Song GG (2012) Associations between TLR polymorphisms and systemic lupus erythematosus: a systematic review and meta-analysis. Clin Exp Rheumatol 30: 262-265.

233. Wang CM, Chang SW, Wu YJ, Lin JC, Ho HH, Chou TC, Yang B, Wu J, Chen JY (2014) Genetic variations in Toll-like receptors (TLRs 3/7/8) are associated with systemic lupus erythematosus in a Taiwanese population. Sci Rep 4: 3792.

234. Sawalha AH, Webb R, Han S, Kelly JA, Kaufman KM, Kimberly RP, Alarcon-Riquelme ME, James JA, Vyse TJ, Gilkeson GS, Choi CB, Scofield RH, Bae SC, Nath SK, Harley JB (2008) Common variants within MECP2 confer risk of systemic lupus erythematosus. PLoS One 3: e1727.

235. Webb R, Wren JD, Jeffries M, Kelly JA, Kaufman KM, Tang Y, Frank MB, Merrill J, Kimberly RP, Edberg JC, Ramsey-Goldman R, Petri M, Reveille JD, Alarcon GS, Vila LM, Alarcon-Riquelme ME, James JA, Vyse TJ, Moser KL, Gaffney PM,

Gilkeson GS, Harley JB, Sawalha AH (2009) Variants within MECP2, a key transcription regulator, are associated with increased susceptibility to lupus and differential gene expression in patients with systemic lupus erythematosus. Arthritis Rheum 60: 1076-1084.

236. Moser KL, Kelly JA, Lessard CJ, Harley JB (2009) Recent insights into the genetic basis of systemic lupus erythematosus. Genes Immun 10: 373-379.

237. Deng Y, Tsao BP (2010) Genetic susceptibility to systemic lupus erythematosus in the genomic era. Nat Rev Rheumatol 6: 683-692.

238. Siontis KC, Patsopoulos NA, Ioannidis JP (2010) Replication of past candidate loci for common diseases and phenotypes in 100 genome-wide association studies. Eur J Hum Genet 18: 832-837.

239. Pe'er I, Yelensky R, Altshuler D, Daly MJ (2008) Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genetic Epidemiology 32: 381-385.

240. Harley JB (2007) IL-7Rα and multiple sclerosis risk. Nature Genetics 39: 1053-1054.

241. Lessard CJ, Adrianto I, Kelly JA, Kaufman KM, Grundahl KM, Adler A, Williams AH, Gallant CJ, Anaya J-M, Bae S-C, Boackle SA, Brown EE, Chang D-M, Criswell LA, Edberg JC, Freedman BI, Gregersen PK, Gilkeson GS, Jacob CO, James JA, Kamen DL, Kimberly RP, Martin J, Merrill JT, Niewold TB, Park S-Y, Petri MA, Pons-Estel BA, Ramsey-Goldman R, Reveille JD, Song YW, Stevens AM, Tsao BP, Vila LM, Vyse TJ, Yu C-Y, Guthridge JM, Bruner GR, Langefeld

CD, Montgomery C, Harley JB, Scofield RH, Gaffney PM, Moser KL, Networks16 MEA-RobotBaG (2011) Identification of a Systemic Lupus Erythematosus Susceptibility Locus at 11p13 between PDHX and CD44 in a Multiethnic Study. The American Journal of Human Genetics 88: 83-91.

242. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, Bakker SF, Bardella MT, Bhaw-Rosun L, Castillejo G, De La Concha EG, De Almeida RC, Dias K-RM, Van Diemen CC, Dubois PCA, Duerr RH, Edkins S, Franke L, Fransen K, Gutierrez J, Heap GAR, Hrdlickova B, Hunt S, Izurieta LP, Izzo V, Joosten LAB, Langford C, Mazzilli MC, Mein CA, Midah V, Mitrovic M, Mora B, Morelli M, Nutland S, Núñez C, Onengut-Gumuscu S, Pearce K, Platteel M, Polanco I, Potter S, Ribes-Koninckx C, Ricaño-Ponce I, Rich SS, Rybak A, Santiago JL, Senapati S, Sood A, Szajewska H, Troncone R, Varadé J, Wallace C, Wolters VM, Zhernakova A, CEGEC SCotGoCD, Group PS, WTCCC WTCCC, Thelma BK, Cukrowska B, Urcelay E, Bilbao JR, Mearin ML, Barisani D, Barrett JC, Plagnol V, Deloukas P, Wijmenga C, Van Heel DA (2011) Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease-supp. Nature Genetics 43: 1193-1201.

243. Nath SK, Han S, Kim-Howard X, Kelly JA, Viswanathan P, Gilkeson GS, Chen W, Zhu C, Mcever RP, Kimberly RP, Alarcón-Riquelme ME, Vyse TJ, Li Q-Z, Wakeland EK, Merrill JT, James JA, Kaufman KM, Guthridge JM, Harley JB (2008) A nonsynonymous functional variant in integrin-alpha(M) (encoded by

ITGAM) is associated with systemic lupus erythematosus. Nature Genetics 40: 152-154.

244. Wen F, Ellingson SM, Kyogoku C, Peterson EJ, Gaffney PM (2011) Exon 6 variants carried on systemic lupus erythematosus (SLE) risk haplotypes modulate IRF5 function. Autoimmunity 44: 82-89.

245. Lofgren SE, Yin H, Delgado-Vega AM, Sanchez E, Lewen S, Pons-Estel BA, Witte T, D'Alfonso S, Ortego-Centeno N, Martin J, Alarcon-Riquelme ME, Kozyrev SV (2010) Promoter insertion/deletion in the IRF5 gene is highly associated with susceptibility to systemic lupus erythematosus in distinct populations, but exerts a modest effect on gene expression in peripheral blood mononuclear cells. J Rheumatol 37: 574-578.

246. Hu X, Kim H, Stahl E, Plenge R, Daly M, Raychaudhuri S (2011) Integrating autoimmune risk Loci with gene-expression data identifies specific pathogenic immune cell subsets. Am J Hum Genet 89: 496-506.

247. Harley ITW, Kaufman KM, Langefeld CD, Harley JB, Kelly JA (2009) Genetic susceptibility to SLE: new insights from fine mapping and genome-wide association studies. Nature Reviews Genetics 10: 285-290.

248. Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. Proceedings of the National Academy of Sciences 109: 1193-1198.

249. Sherer Y, Gorstein A, Fritzler MJ, Shoenfeld Y (2004) Autoantibody explosion in systemic lupus erythematosus: More than 100 different antibodies found in SLE patients. Seminars in Arthritis and Rheumatism 34: 501-537.

250. Tsokos GC (2011) Systemic lupus erythematosus. New England Journal of Medicine 365: 2110-2121.

251. Cervera R, Piette JC, Font J, Khamashta MA, Shoenfeld Y, Camps MT, Jacobsen S, Lakos G, Tincani A, Kontopoulou-Griva I, Galeazzi M, Meroni PL, Derksen RH, de Groot PG, Gromnica-Ihle E, Baleva M, Mosca M, Bombardieri S, Houssiau F, Gris JC, Quere I, Hachulla E, Vasconcelos C, Roch B, Fernandez-Nebro A, Boffa MC, Hughes GR, Ingelmo M (2002) Antiphospholipid syndrome: clinical and immunologic manifestations and patterns of disease expression in a cohort of 1,000 patients. Arthritis and Rheumatism 46: 1019-1027.

252. Dal Porto JM, Gauld SB, Merrell KT, Mills D, Pugh-Bernard AE, Cambier J (2004) B cell antigen receptor signaling 101. Mol Immunol 41: 599-613.

253. Kurosaki T, Shinohara H, Baba Y (2010) B Cell Signaling and Fate Decision. Annual Review of Immunology 28: 21-55.

254. Hou P, Araujo E, Zhao T, Zhang M, Massenburg D, Veselits M, Doyle C, Dinner AR, Clark MR (2006) B Cell Antigen Receptor Signaling and Internalization Are Mutually Exclusive Events. PLoS Biology 4: e200.

255. Pillai S, Mattoo H, Cariappa A (2011) B cells and autoimmunity. Current Opinion in Immunology 23: 721-731.

256. Lu R, Vidal GS, Kelly JA, Delgado-Vega AM, Howard XK, Macwana SR, Dominguez N, Klein W, Burrell C, Harley IT, Kaufman KM, Bruner GR, Moser KL, Gaffney PM, Gilkeson GS, Wakeland EK, Li QZ, Langefeld CD, Marion MC, Divers J, Alarcon GS, Brown EE, Kimberly RP, Edberg JC, Ramsey-Goldman R, Reveille JD, McGwin G, Jr., Vila LM, Petri MA, Bae SC, Cho SK, Bang SY, Kim I, Choi CB, Martin J, Vyse TJ, Merrill JT, Harley JB, Alarcon-Riquelme ME, Nath SK, James JA, Guthridge JM (2009) Genetic associations of LYN with systemic lupus erythematosus. Genes and Immunity 10: 397-403.

257. Flores-Borja F, Kabouridis PS, Jury EC, Isenberg DA, Mageed RA (2005) Decreased Lyn expression and translocation to lipid raft signaling domains in B lymphocytes from patients with systemic lupus erythematosus. Arthritis and Rheumatism 52: 3955-3965.

258. Flores-Borja F, Kabouridis PS, Jury EC, Isenberg DA, Mageed RA (2006) Altered lipid raft–associated proximal signaling and translocation of CD45 tyrosine phosphatase in B lymphocytes from patients with systemic lupus erythematosus. Arthritis and Rheumatism 56: 291-302.

259. Nishizumi H, Taniuchi I, Yamanashi Y, Kitamura D, Ilic D, Mori S, Watanabe T, Yamamoto T (1995) Impaired proliferation of peripheral B cells and indication of autoimmune disease in lyn-deficient mice. Immunity 3: 549-560.

260. Hibbs ML, Tarlinton DM, Armes J, Grail D, Hodgson G, Maglitto R, Stacker SA, Dunn AR (1995) Multiple defects in the immune system of Lyn-deficient mice, culminating in autoimmune disease. Cell 83: 301-311.

261. Saijo K, Schmedt C, Su I-h, Karasuyama H, Lowell CA, Reth M, Adachi T, Patke A, Santana A, Tarakhovsky A (2003) Essential role of Src-family protein tyrosine kinases in NF-κB activation during B cell development. Nature Immunology 4: 274-279.

262. Chan VW, Meng F, Soriano P, DeFranco AL, Lowell CA (1997) Characterization of the B lymphocyte populations in Lyn-deficient mice and the role of Lyn in signal initiation and down-regulation. Immunity 7: 69-81.

263. Wang J, Koizumi T, Watanabe T (1996) Altered antigen receptor signaling and impaired Fas-mediated apoptosis of B cells in Lyn-deficient mice. The Journal of experimental medicine 184: 831-838.

264. Gross AJ, Proekt I, DeFranco AL (2011) Elevated BCR signaling and decreased survival of Lyn-deficient transitional and follicular B cells. European Journal of Immunology 41: 3645-3655.

265. Gross AJ, Lyandres JR, Panigrahi AK, Prak ETL, DeFranco AL (2009) Developmental Acquisition of the Lyn-CD22-SHP-1 Inhibitory Pathway Promotes B Cell Tolerance. The Journal of Immunology 182: 5382-5392.

266. Cornall RJ, Cyster JG, Hibbs ML, Dunn AR, Otipoby KL, Clark EA, Goodnow CC (1998) Polygenic autoimmune traits: Lyn, CD22, and SHP-1 are limiting elements of a biochemical pathway regulating BCR signaling and selection. Immunity 8: 497-508.

267. Scapini P, Hu Y, Chu C-L, Migone T-S, Defranco AL, Cassatella MA, Lowell CA (2010) Myeloid cells, BAFF, and IFN- establish an inflammatory loop that

exacerbates autoimmunity in Lyn-deficient mice. Journal of Experimental Medicine 207: 1757-1773.

268. Silver KL, Crockford TL, Bouriez-Jones T, Milling S, Lambe T, Cornall RJ (2007) MyD88-dependent autoimmune disease in Lyn-deficient mice. European Journal of Immunology 37: 2734-2743.

269. Dymecki SM, Niederhuber JE, Desiderio SV (1990) Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells. Science 247: 332-336.

270. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA, Ferreira RC, Pant PVK, Ballinger DG, Kosoy R, Demirci FY, Kamboh MI, Kao AH, Tian C, Gunnarsson I, Bengtsson AA, Rantapää-Dahlqvist S, Petri M, Manzi S, Seldin MF, Rönnblom L, Syvänen A-C, Criswell LA, Gregersen PK, Behrens TW (2008) Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. The New England Journal of Medicine 358: 900-909.

271. Ito I, Kawasaki A, Ito S, Hayashi T, Goto D, Matsumoto I, Tsutsumi A, Hom G, Graham RR, Takasaki Y, Hashimoto H, Ohashi J, Behrens TW, Sumida T, Tsuchiya N (2009) Replication of the association between the C8orf13-BLKregion and systemic lupus erythematosus in a Japanese population. Arthritis and Rheumatism 60: 553-558.

272. Castillejo-López C, Delgado-Vega AM, Wojcik J, Kozyrev SV, Thavathiru E, Wu YY, Sanchez E, Pollmann D, Lopez-Egido JR, Fineschi S, Dominguez N, Lu R, James JA, Merrill JT, Kelly JA, Kaufman KM, Moser KL, Gilkeson G, Frostegard J, Pons-Estel BA, D'alfonso S, Witte T, Callejas JL, Harley JB, Gaffney PM,

Martin J, Guthridge JM, Alarcon-Riquelme ME (2011) Genetic and physical interaction of the B-cell systemic lupus erythematosus-associated genes BANK1 and BLK. Annals of the Rheumatic Diseases 71: 136-142.

273. Texido G, Su IH, Mecklenbräuker I, Saijo K, Malek SN, Desiderio S, Rajewsky K, Tarakhovsky A (2000) The B-cell-specific Src-family kinase Blk is dispensable for B-cell development and activation. Molecular and Cellular Biology 20: 1227-1233.

274. Maas A, Hendriks RW (2001) Role of Bruton's tyrosine kinase in B cell development. Developmental Immunology 8: 171-181.

275. Yokoyama K, Su Ih IH, Tezuka T, Yasuda T, Mikoshiba K, Tarakhovsky A, Yamamoto T (2002) BANK regulates BCR-induced calcium mobilization by promoting tyrosine phosphorylation of IP(3) receptor. EMBO J 21: 83-92.

276. Aiba Y, Yamazaki T, Okada T, Gotoh K, Sanjo H, Ogata M, Kurosaki T (2006) BANK Negatively Regulates Akt Activation and Subsequent B Cell Responses. Immunity 24: 259-268.

277. Kozyrev SV, Abelson A-K, Wojcik J, Zaghlool A, Linga Reddy MVP, Sanchez E, Gunnarsson I, Svenungsson E, Sturfelt G, Jönsen A, Truedsson L, Pons-Estel BA, Witte T, D'alfonso S, Barrizzone N, Danieli MG, Gutierrez C, Suarez A, Junker P, Laustrup H, González-Escribano MF, Martin J, Abderrahim H, Alarcón-Riquelme ME (2008) Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nature Genetics 40: 211-216.

278. Guo L, Deshmukh H, Lu R, Vidal GS, Kelly JA, Kaufman KM, Dominguez N, Klein W, Kim-Howard X, Bruner GR, Scofield RH, Moser KL, Gaffney PM, Dozmorov

IM, Gilkeson GS, Wakeland EK, Li Q-Z, Langefeld CD, Marion MC, Williams AH, Divers J, n GSAo, Brown EE, Kimberly RP, Edberg JC, Ramsey-Goldman R, Reveille JD, McGwin G, aacute LMV, Petri MA, Vyse TJ, Merrill JT, James JA, Nath SK, Harley JB, Guthridge JM (2009) Replication of the BANK1 genetic association with systemic lupus erythematosus in a European-derived population. Genes and Immunity 10: 531-538.

279. Kozyrev SV, Bernal-Quiros M, Alarcon-Riquelme ME, Castillejo-Lopez C (2012) The dual effect of the lupus-associated polymorphism rs10516487 on BANK1 gene expression and protein localization. Genes Immun 13: 129-138.

280. Cohen S, Dadi H, Shaoul E, Sharfe N, Roifman CM (1999) Cloning and characterization of a lymphoid-specific, inducible human protein tyrosine phosphatase, Lyp. Blood 93: 2013-2024.

281. Begovich AB, Carlton VEH, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, Huang Q, Smith AM, Spoerke JM, Conn MT, Chang M, Chang S-YP, Saiki RK, Catanese JJ, Leong DU, Garcia VE, McAllister LB, Jeffery DA, Lee AT, Batliwalla F, Remmers E, Criswell LA, Seldin MF, Kastner DL, Amos CI, Sninsky JJ, Gregersen PK (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 75: 330-337.

282. Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VEH, Chang M, Ramos P, Baechler EC, Batliwalla FM, Novitzke J, Williams AH, Gillett C, Rodine P, Graham RR, Ardlie KG, Gaffney PM, Moser KL, Petri M, Begovich AB,

Gregersen PK, Behrens TW (2004) Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. American journal of human genetics 75: 504-507.

283. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, MacMurray J, Meloni GF, Lucarelli P, Pellecchia M, Eisenbarth GS, Comings D, Mustelin T (2004) A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nature Genetics 36: 337-338.

284. Smyth D, Cooper JD, Collins JE, Heward JM, Franklyn JA, Howson JM, Vella A, Nutland S, Rance HE, Maier L, Barratt BJ, Guja C, Ionescu-Tirgoviste C, Savage DA, Dunger DB, Widmer B, Strachan DP, Ring SM, Walker N, Clayton DG, Twells RC, Gough SC, Todd JA (2004) Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes, and evidence for its role as a general autoimmunity locus. Diabetes 53: 3020-3023.

285. Vandiedonck C, Capdevielle C, Giraud M, Krumeich S, Jais JP, Eymard B, Tranchant C, Gajdos P, Garchon HJ (2006) Association of the PTPN22*R620W polymorphism with autoimmune myasthenia gravis. Annals of Neurology 59: 404-407.

286. Lea W, Lee Y (2011) The association between the PTPN22 C1858T polymorphism and systemic lupus erythematosus: a meta-analysis update. Lupus 20: 51-57.

287. Vang T, Congia M, Macis MD, Musumeci L, Orru V, Zavattari P, Nika K, Tautz L, Tasken K, Cucca F, Mustelin T, Bottini N (2005) Autoimmune-associated

lymphoid tyrosine phosphatase is a gain-of-function variant. Nat Genet 37: 1317-1319.

288. Fiorillo E, Orru V, Stanford SM, Liu Y, Salek M, Rapini N, Schenone AD, Saccucci P, Delogu LG, Angelini F, Manca Bitti ML, Schmedt C, Chan AC, Acuto O, Bottini N (2010) Autoimmune-associated PTPN22 R620W Variation Reduces Phosphorylation of Lymphoid Phosphatase on an Inhibitory Tyrosine Residue. Journal of Biological Chemistry 285: 26506-26518.

289. Rieck M, Arechiga A, Onengut-Gumuscu S, Greenbaum C, Concannon P, Buckner JH (2007) Genetic variation in PTPN22 corresponds to altered function of T and B lymphocytes. The Journal of Immunology 179: 4704-4710.

290. Lefvert AK, Zhao Y, Ramanujam R, Yu S, Pirskanen R, Hammarström L (2008) PTPN22 R620W promotes production of anti-AChR autoantibodies and IL-2 in myasthenia gravis. Journal of Neuroimmunology 197: 110-113.

291. Zhang J, Zahir N, Jiang Q, Miliotis H, Heyraud S, Meng X, Dong B, Xie G, Qiu F, Hao Z, McCulloch CA, Keystone EC, Peterson AC, Siminovitch KA (2011) The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. Nat Genet 43: 902-907.

292. Habib T, Funk A, Rieck M, Brahmandam A, Dai X, Panigrahi AK, Luning Prak ET, Meyer-Bahlburg A, Sanda S, Greenbaum C, Rawlings DJ, Buckner JH (2011) Altered B Cell Homeostasis Is Associated with Type I Diabetes and Carriers of the PTPN22 Allelic Variant. The Journal of Immunology 188: 487-496.

293. Menard L, Saadoun D, Isnardi I, Ng Y-S, Meyers G, Massad C, Price C, Abraham C, Motaghedi R, Buckner JH, Gregersen PK, Meffre E (2011) The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. Journal of Clinical Investigation 121: 3635-3644.

294. Arechiga AF, Habib T, He Y, Zhang X, Zhang ZY, Funk A, Buckner JH (2009) Cutting Edge: The PTPN22 Allelic Variant Associated with Autoimmunity Impairs B Cell Signaling. The Journal of Immunology 182: 3343-3347.

295. Salloum R, Niewold TB (2011) Interferon regulatory factors in human lupus pathogenesis. Translational Research 157: 326-331.

296. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M, Parkin M, Barry R, Winslow W, Healy C, Graham RR, Neale BM, Izmailova E, Roubenoff R, Parker AN, Glass R, Karlson EW, Maher N, Hafler DA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Coblyn J, Weinblatt ME, Gabriel SB, Purcell S, Klareskog L, Gregersen PK, Shadick NA, Daly MJ, Altshuler D (2007) Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nature Genetics 39: 1477-1482.

297. Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, Wilson AG, Marinou I, Morgan A, Emery P, Carter A, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Strachan D, Worthington J (2007) Rheumatoid arthritis association at 6q23. Nature Genetics 39: 1431-1433.

298. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burtt NP, Guiducci C, Parkin M, Gates C, Plenge RM, Behrens TW, Wither JE, Rioux JD, Fortin PR, Cunninghame Graham D, Wong AK, Vyse TJ, Daly MJ, Altshuler D, Moser KL, Gaffney PM (2008) Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nature Genetics 40: 1059-1061.

299. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Ilyas Kamboh M, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok P-Y, Criswell LA (2008) Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nature Genetics 40: 1062-1064.

300. Trynka G, Zhernakova A, Romanos J, Franke L, Hunt KA, Turner G, Bruinenberg M, Heap GA, Platteel M, Ryan AW, de Kovel C, Holmes GK, Howdle PD, Walters JR, Sanders DS, Mulder CJ, Mearin ML, Verbeek WH, Trimble V, Stevens FM, Kelleher D, Barisani D, Bardella MT, McManus R, van Heel DA, Wijmenga C (2009) Coeliac disease-associated risk variants in TNFAIP3 and REL implicate altered NF-kappaB signalling. Gut 58: 1078-1083.

301. Wang K, Baldassano R, Zhang H, Qu HQ, Imielinski M, Kugathasan S, Annese V, Dubinsky M, Rotter JI, Russell RK, Bradfield JP, Sleiman PM, Glessner JT, Walters T, Hou C, Kim C, Frackelton EC, Garris M, Doran J, Romano C, Catassi C, Van Limbergen J, Guthery SL, Denson L, Piccoli D, Silverberg MS, Stanley CA, Monos D, Wilson DC, Griffiths A, Grant SF, Satsangi J, Polychronakos C, Hakonarson H (2010) Comparative genetic analysis of inflammatory bowel

disease and type 1 diabetes implicates multiple loci with opposite effects. Human Molecular Genetics 19: 2059-2067.

302. Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H, Wicker LS, Todd JA (2009) Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. Genes and Immunity 10: 188-191.

303. Duffin KC, Woodcock J, Krueger GG (2010) Genetic variations associated with psoriasis and psoriatic arthritis found by genome-wide association. Dermatol Ther 23: 101-113.

304. Adrianto I, Wen F, Templeton A, Wiley G, King JB, Lessard CJ, Bates JS, Hu Y, Kelly JA, Kaufman KM, Guthridge JM, Alarcón-Riquelme ME, Networks BaG, Anaya J-M, Bae S-C, Bang S-Y, Boackle SA, Brown EE, Petri MA, Gallant C, Ramsey-Goldman R, Reveille JD, Vila LM, Criswell LA, Edberg JC, Freedman BI, Gregersen PK, Gilkeson GS, Jacob CO, James JA, Kamen DL, Kimberly RP, Martin J, Merrill JT, Niewold TB, Park S-Y, Pons-Estel BA, Scofield RH, Stevens AM, Tsao BP, Vyse TJ, Langefeld CD, Harley JB, Moser KL, Webb CF, Humphrey MB, Montgomery CG, Gaffney PM (2011) Association of a functional variant downstream of TNFAIP3 with systemic lupus erythematosus. Nature Genetics 43: 253-258.

305. Sarma V, Lin Z, Clark L, Rust BM, Tewari M, Noelle RJ, Dixit VM (1995) Activation of the B-cell surface receptor CD40 induces A20, a novel zinc finger protein that inhibits apoptosis. J Biol Chem 270: 12343-12346.

306. Hymowitz SG, Wertz IE (2010) A20: from ubiquitin editing to tumour suppression. Nat Rev Cancer 10: 1-9.

307. Skaug B, Chen J, Du F, He J, Ma A, Chen ZJ (2011) Direct, noncatalytic mechanism of IKK inhibition by A20. Molecular Cell 44: 559-571.

308. Lee EG (2000) Failure to Regulate TNF-Induced NF-kappa B and Cell Death Responses in A20-Deficient Mice. Science 289: 2350-2354.

309. Turer EE, Tavares RM, Mortier E, Hitotsumatsu O, Advincula R, Lee B, Shifrin N, Malynn BA, Ma A (2008) Homeostatic MyD88-dependent signals cause lethal inflamMation in the absence of A20. Journal of Experimental Medicine 205: 451-464.

310. Harhaj EW, Dixit VM (2012) Regulation of NF-κB by deubiquitinases. Immunological reviews 246: 107-124.

311. Chu Y, Vahl JC, Kumar D, Heger K, Bertossi A, Wojtowicz E, Soberon V, Schenten D, Mack B, Reutelshofer M, Beyaert R, Amann K, van Loo G, Schmidt-Supprian M (2011) B cells lacking the tumor suppressor TNFAIP3/A20 display impaired differentiation and hyperactivation and cause inflammation and autoimmunity in aged mice. Blood 117: 2227-2236.

312. Tavares RM, Turer EE, Liu CL, Advincula R, Scapini P, Rhee L, Barrera J, Lowell CA, Utz PJ, Malynn BA, Ma A (2010) The Ubiquitin Modifying Enzyme A20 Restricts B Cell Survival and Prevents Autoimmunity. Immunity 33: 181-191.

313. Witte T (2007) IgM Antibodies Against dsDNA in SLE. Clinical Reviews in Allergy and Immunology 34: 345-347.

314. Hao Z, Duncan GS, Seagal J, Su Y-W, Hong C, Haight J, Chen N-J, Elia A, Wakeham A, Li WY, Liepa J, Wood GA, Casola S, Rajewsky K, Mak TW (2008) Fas Receptor Expression in Germinal-Center B Cells Is Essential for T and B Lymphocyte Homeostasis. Immunity 29: 615-627.

315. Han J-W, Zheng H-F, Cui Y, Sun L-D, Ye D-Q, Hu Z, Xu J-H, Cai Z-M, Huang W, Zhao G-P, Xie H-F, Fang H, Lu Q-J, Xu J-H, Li X-P, Pan Y-F, Deng D-Q, Zeng F-Q, Ye Z-Z, Zhang X-Y, Wang Q-W, Hao F, Ma L, Zuo X-B, Zhou F-S, Du W-H, Cheng Y-L, Yang J-Q, Shen S-K, Li J, Sheng Y-J, Zuo X-X, Zhu W-F, Gao F, Zhang P-L, Guo Q, Li B, Gao M, Xiao F-L, Quan C, Zhang C, Zhang Z, Zhu K-J, Li Y, Hu D-Y, Lu W-S, Huang J-L, Liu S-X, Li H, Ren Y-Q, Wang Z-X, Yang C-J, Wang P-G, Zhou W-M, Lv Y-M, Zhang A-P, Zhang S-Q, Lin D, Li Y, Low HQ, Shen M, Zhai Z-F, Wang Y, Zhang F-Y, Yang S, Liu J-J, Zhang X-J (2009) Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nature Genetics 41: 1234-1237.

316. Shiote Y, Ouchida M, Jitsumori Y, Ogama Y, Matsuo Y, Ishimaru F, Tanimoto M, Shimizu K (2006) Multiple splicing variants of Naf1/ABIN-1 transcripts and their alterations in hematopoietic tumors. International journal of molecular medicine 18: 917-924.

317. Heyninck K, Kreike MM, Beyaert R (2003) Structure-function analysis of the A20-binding inhibitor of NF-kappa B activation, ABIN-1. FEBS Lett 536: 135-140.

318. Ashida H, Kim M, Schmidt-Supprian M, Ma A, Ogawa M, Sasakawa C (2009) A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKγ to dampen the host NF-κB-mediated inflammatory response. Nature Cell Biology 12: 66-73.

319. Zhang S, Fukushi M, Hashimoto S, Gao C, Huang L, Fukuyo Y, Nakajima T, Amagasa T, Enomoto S, Koike K, Miura O, Yamamoto N, Tsuchida N (2002) A new ERK2 binding protein, Naf1, attenuates the EGF/ERK2 nuclear signaling. Biochemical and Biophysical Research Communications 297: 17-23.

320. Gao L, Coope H, Grant S, Ma A, Ley SC, Harhaj EW (2011) ABIN1 Protein Cooperates with TAX1BP1 and A20 Proteins to Inhibit Antiviral Signaling. Journal of Biological Chemistry 286: 36592-36602.

321. Chen F, Bhatia D, Chang Q, Castranova V (2006) Finding NEMO by K63-linked polyubiquitin chain. Cell Death and Differentiation 13: 1835-1838.

322. Zhou J, Wu R, High AA, Slaughter CA, Finkelstein D, Rehg JE, Redecke V, Häcker H (2011) A20-binding inhibitor of NF-κB (ABIN1) controls Toll-like receptor-mediated CCAAT/enhancer-binding protein β activation and protects from inflammatory disease. Proceedings of the National Academy of Sciences 108: E998-E1006.

323. Nanda SK, Venigalla RKC, Ordureau A, Patterson-Kane JC, Powell DW, Toth R, C Arthur JS, Cohen P (2011) Polyubiquitin binding to ABIN1 is required to prevent autoimmunity. Journal of Experimental Medicine 208: 1215-1228.

324. Namjou B, Choi CB, Harley IT, Alarcon-Riquelme ME, Kelly JA, Glenn SB, Ojwang JO, Adler A, Kim K, Gallant CJ, Boackle SA, Criswell LA, Kimberly RP, Brown

EE, Edberg J, Alarcon GS, Stevens AM, Jacob CO, Gilkeson GS, Kamen DL, Tsao BP, Anaya JM, Kim EM, Park SY, Sung YK, Guthridge JM, Merrill JT, Petri M, Ramsey-Goldman R, Vila LM, Niewold TB, Martin J, Pons-Estel BA, Vyse TJ, Freedman BI, Moser KL, Gaffney PM, Williams AH, Comeau ME, Reveille JD, Kang C, James JA, Scofield RH, Langefeld CD, Kaufman KM, Harley JB, Bae SC (2012) Evaluation of TRAF6 in a large multi-ancestral lupus cohort. Arthritis Rheum.

325. Hildebrand JM, Luo Z, Manske MK, Price-Troska T, Ziesmer SC, Lin W, Hostager BS, Slager SL, Witzig TE, Ansell SM, Cerhan JR, Bishop GA, Novak AJ (2010) A BAFF-R mutation associated with non-Hodgkin lymphoma alters TRAF recruitment and reveals new insights into BAFF-R signaling. Journal of Experimental Medicine 207: 2569-2579.

326. Inoue J, Ishida T, Tsukamoto N, Kobayashi N, Naito A, Azuma S, Yamamoto T (2000) Tumor necrosis factor receptor-associated factor (TRAF) family: adapter proteins that mediate cytokine signaling. Experimental Cell Research 254: 14-24.

327. Anolik JH, Barnard J, Cappione A, Pugh-Bernard AE, Felgar RE, Looney RJ, Sanz Ia (2004) Rituximab improves peripheral B cell abnormalities in human systemic lupus erythematosus. Arthritis & Rheumatism 50: 3580-3590.

328. Leandro MJ, Edwards JC, Cambridge G, Ehrenstein MR, Isenberg DA (2002) An open study of B lymphocyte depletion in systemic lupus erythematosus. Arthritis and Rheumatism 46: 2673-2677.

329. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, Utset TO, Gordon C, Isenberg DA, Hsieh H-J, Zhang D, Brunetta PG (2010) Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: The randomized, double-blind, phase ii/iii systemic lupus erythematosus evaluation of rituximab trial. Arthritis & Rheumatism 62: 222-233.

330. Furie R, Petri M, Zamani O, Cervera R, Wallace DJ, Tegzova D, Sanchez-Guerrero J, Schwarting A, Merrill JT, Chatham WW, Stohl W, Ginzler EM, Hough DR, Zhong ZJ, Freimuth W, van Vollenhoven RF (2011) A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. Arthritis and Rheumatism 63: 3918-3930.

331. McClain MT, Heinlen LD, Dennis GJ, Roebuck J, Harley JB, James JA (2004) Early events in lupus humoral autoimmunity suggest initiation through molecular mimicry. Nature Medicine 11: 85-89.

332. James JA, Neas BR, Moser KL, Hall T, Bruner GR, Sestak AL, Harley JB (2001) Systemic lupus erythematosus in adults is associated with previous Epstein-Barr virus exposure. Arthritis and Rheumatism 44: 1122-1126.

333. Moon UY, Park SJ, Oh ST, Kim W-U, Park S-H, Lee S-H, Cho C-S, Kim H-Y, Lee W-K, Lee SK (2004) Patients with systemic lupus erythematosus have abnormally elevated Epstein-Barr virus load in blood. Arthritis Research and Therapy 6: R295-302.

334. Dykstra ML, Longnecker R, Pierce SK (2001) Epstein-Barr virus coopts lipid rafts to block the signaling and antigen transport functions of the BCR. Immunity 14: 57-67.

335. Ning S, Pagano JS (2010) The A20 Deubiquitinase Activity Negatively Regulates LMP1 Activation of IRF7. Journal of Virology 84: 6130-6138.

336. Clark MR, Tanaka A, Powers SE, Veselits M (2011) Receptors, subcellular compartments and the regulation of peripheral B cell responses: the illuminating state of anergy. Molecular Immunology 48: 1281-1286.

337. Lander ES (2011) Initial impact of the sequencing of the human genome. Nature 470: 187-197.

338. Rasmussen A, Sevier S, Kelly JA, Glenn SB, Aberle T, Cooney CM, Grether A, James E, Ning J, Tesiram J, Morrisey J, Powe T, Drexel M, Daniel W, Namjou B, Ojwang JO, Nguyen KL, Cavett JW, Te JL, James JA, Scofield RH, Moser K, Gilkeson GS, Kamen DL, Carson CW, Quintero-del-Rio AI, del Carmen Ballesteros M, Punaro MG, Karp DR, Wallace DJ, Weisman M, Merrill JT, Rivera R, Petri MA, Albert DA, Espinoza LR, Utset TO, Shaver TS, Arthur E, Anaya JM, Bruner GR, Harley JB (2011) The lupus family registry and repository. Rheumatology (Oxford) 50: 47-59.

339. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics 81: 559-575.

340. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39: 906-913.

341. Marchini J, Howie B (2010) Genotype imputation for genome-wide association studies. Nat Rev Genet 11: 499-511.

342. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-678.

343. Wellcome Trust Case Control C, Maller JB, McVean G, Byrnes J, Vukcevic D, Palin K, Su Z, Howson JM, Auton A, Myers S, Morris A, Pirinen M, Brown MA, Burton PR, Caulfield MJ, Compston A, Farrall M, Hall AS, Hattersley AT, Hill AV, Mathew CG, Pembrey M, Satsangi J, Stratton MR, Worthington J, Craddock N, Hurles M, Ouwehand W, Parkes M, Rahman N, Duncanson A, Todd JA, Kwiatkowski DP, Samani NJ, Gough SC, McCarthy MI, Deloukas P, Donnelly P (2012) Bayesian refinement of association signals for 14 loci in 3 common diseases. Nat Genet 44: 1294-1301.

344. Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Bonnen PE, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Gonzaga-Jauregui C, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Zhang Q, Ghori MJ, McGinnis R, McLaren

W, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE (2010) Integrating common and rare genetic variation in diverse human populations. Nature 467: 52-58.

345. Gabriel SB (2002) The Structure of Haplotype Blocks in the Human Genome. Science 296: 2225-2229.

346. Molineros JE, Kim-Howard X, Deshmukh H, Jacob CO, Harley JB, Nath SK (2009) Admixture in Hispanic Americans: its impact on ITGAM association and implications for admixture mapping in SLE. Genes and Immunity 10: 539-545.

347. Lessard CJ, Adrianto I, Ice JA, Wiley GB, Kelly JA, Glenn SB, Adler AJ, Li H, Rasmussen A, Williams AH, Ziegler J, Comeau ME, Marion M, Wakeland BE, Liang C, Ramos PS, Grundahl KM, Gallant CJ, Alarcón GS, Anaya J-M, Bae S-C, Boackle SA, Brown EE, Chang D-M, Cho S-K, Criswell LA, Edberg JC, Freedman BI, Gilkeson GS, Jacob CO, James JA, Kamen DL, Kimberly RP, Kim J-H, Martin J, Merrill JT, Niewold TB, Park S-Y, Petri MA, Pons-Estel BA, Ramsey-Goldman R, Reveille JD, Scofield RH, Song YW, Stevens AM, Tsao BP, Vila LM, Vyse TJ, Yu C-Y, Guthridge JM, Kaufman KM, Harley JB, Wakeland EK, Langefeld CD, Gaffney PM, Montgomery CG, Moser KL, Networks17 MEA-RftBaG (2012) Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as Susceptibility Loci for Systemic Lupus Erythematosus in a Large-Scale Multiracial Replication Study. The American Journal of Human Genetics 90: 648-660.

348. Seet L-F, Hong W (2006) The Phox (PX) domain proteins and membrane traffic. Biochim Biophys Acta 1761: 878-896.

349. Mao H, Ferguson TS, Cibulsky SM, Holmqvist M, Ding C, Fei H, Levitan IB (2005) MONaKA, a novel modulator of the plasma membrane Na,K-ATPase. The Journal of neuroscience : the official journal of the Society for Neuroscience 25: 7934-7943.

350. Pedersen NM, Raiborg C, Brech A, Skarpen E, Roxrud I, Platta HW, Liestøl K, Stenmark H (2012) The PtdIns3P-Binding Protein Phafin 2 Mediates Epidermal Growth Factor Receptor Degradation by Promoting Endosome Fusion. Traffic 13: 1547-1563.

351. Grimaldi CM, Hicks R, Diamond B (2005) B cell selection and susceptibility to autoimmunity. Journal of immunology (Baltimore, Md : 1950) 174: 1775-1781.

352. Shlomchik MJ, Madaio MP, Ni D, Trounstein M, Huszar D (1994) The role of B cells in lpr/lpr-induced autoimmunity. J Exp Med 180: 1295-1306.

353. Chan O, Shlomchik MJ (1998) A new role for B cells in systemic autoimmunity: B cells promote spontaneous T cell activation in MRL-lpr/lpr mice. Journal of immunology (Baltimore, Md : 1950) 160: 51-59.

354. Chan OT, Hannum LG, Haberman AM, Madaio MP, Shlomchik MJ (1999) A novel mouse with B cells but lacking serum antibody reveals an antibody-independent role for B cells in murine lupus. The Journal of experimental medicine 189: 1639-1648.

355. Chaturvedi A, Martz R, Dorward D, Waisberg M, Pierce SK (2011) Endocytosed BCRs sequentially regulate MAPK and Akt signaling pathways from intracellular compartments. Nature Immunology 12: 1119-1126.

356. Heesters BA, Das A, Chatterjee P, Carroll MC (2014) Do follicular dendritic cells regulate lupus-specific B cells? Mol Immunol 62: 283-288.

357. Chaturvedi A, Dorward D, Pierce SK (2008) The B Cell Receptor Governs the Subcellular Location of Toll-like Receptor 9 Leading to Hyperresponses to DNA-Containing Antigens. Immunity 28: 799-809.

358. Giltiay NV, Chappell CP, Clark EA (2012) B-cell selection and the development of autoantibodies. Arthritis Res Ther 14 Suppl 4: S1.

359. McKeigue PM, Carpenter JR, Parra EJ, Shriver MD (2000) Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. Ann Hum Genet 64: 171-186.

360. Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, McKeigue PM (2003) Control of confounding of genetic associations in stratified populations. Am J Hum Genet 72: 1492-1504.

361. Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, Kessing BD, Malasky MJ, Scafe C, Le E, De Jager PL, Mignault AA, Yi Z, De The G, Essex M, Sankale JL, Moore JH, Poku K, Phair JP, Goedert JJ, Vlahov D, Williams SM, Tishkoff SA, Winkler CA, De La Vega FM, Woodage T, Sninsky JJ, Hafler DA, Altshuler D, Gilbert DA, O'Brien SJ, Reich D (2004) A

high-density admixture map for disease gene discovery in african americans. Am J Hum Genet 74: 1001-1013.

362. Kottyan LC, Davis BP, Sherrill JD, Liu K, Rochman M, Kaufman K, Weirauch MT, Vaughn S, Lazaro S, Rupert AM, Kohram M, Stucke EM, Kemme KA, Magnusen A, He H, Dexheimer P, Chehade M, Wood RA, Pesek RD, Vickery BP, Fleischer DM, Lindbad R, Sampson HA, Mukkada VA, Putnam PE, Abonia JP, Martin LJ, Harley JB, Rothenberg ME (2014) Genome-wide association analysis of eosinophilic esophagitis provides insight into the tissue specificity of this allergic disease. Nat Genet 46: 895-900.

363. Barrett JC (2009) Haploview: Visualization and analysis of SNP genotype data. Cold Spring Harb Protoc 2009: pdb ip71.

364. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263-265.

365. Sankararaman S, Sridhar S, Kimmel G, Halperin E (2008) Estimating Local Ancestry in Admixed Populations. The American Journal of Human Genetics 82: 290-303.

366. Stephens M, Balding DJ (2009) Modelling: Bayesian statistical methods for genetic association studies. Nature Publishing Group 10: 681-690.

367. Sung YJ, Wang L, Rankinen T, Bouchard C, Rao DC (2012) Performance of genotype imputations using data from the 1000 Genomes Project. Hum Hered 73: 18-25.

368. Chen J, Zhang JG, Li J, Pei YF, Deng HW (2013) On combining reference data to improve imputation accuracy. PLoS One 8: e55600.

369. Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19: 149-150.

370. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

371. Malhotra S, Kovats S, Zhang W, Coggeshall KM (2009) B cell antigen receptor endocytosis and antigen presentation to T cells require Vav and dynamin. Journal of Biological Chemistry 284: 24088-24097.

372. Malhotra S, Kovats S, Zhang W, Coggeshall KM (2009) Vav and Rac activation in B cell antigen receptor endocytosis involves Vav recruitment to the adapter protein LAB. Journal of Biological Chemistry 284: 36202-36212.

373. Howie B, Marchini J, Stephens M (2011) Genotype imputation with thousands of genomes. G3: Genesl Genomesl Genetics 1: 457-470.

374. Ramos PS, Williams AH, Ziegler JT, Comeau ME, Guy RT, Lessard CJ, Li H, Edberg JC, Zidovetzki R, Criswell LA, Gaffney PM, Graham DC, Graham RR, Kelly JA, Kaufman KM, Brown EE, Alarcón GS, Petri MA, Reveille JD, McGwin G, Vila LM, Ramsey-Goldman R, Jacob CO, Vyse TJ, Tsao BP, Harley JB, Kimberly RP, Alarcón-Riquelme ME, Langefeld CD, Moser KL (2011) Genetic

analyses of interferon pathway-related genes reveal multiple new loci associated with systemic lupus erythematosus. Arthritis & Rheumatism 63: 2049-2057.

375. Al-Mayouf SM, Sunker A, Abdwani R, Abrawi SA, Almurshedi F, Alhashmi N, Al Sonbul A, Sewairi W, Qari A, Abdallah E, Al-Owain M, Al Motywee S, Al-Rayes H, Hashem M, Khalak H, Al-Jebali L, Alkuraya FS (2011) Loss-of-function variant in DNASE1L3 causes a familial form of systemic lupus erythematosus. Nat Genet 43: 1186-1188.

376. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV, Assassi S, Ying J, Tan FK, Arnett FC, Reveille JD, Guerra S, Teruel M, Carmona FD, Gregersen PK, Lee AT, López-Isac E, Ochoa E, Carreira P, Simeón CP, Castellví I, González-Gay MÁ, Group tSS, Ortego-Centeno N, Ríos R, Callejas JL, Navarrete N, Portales RG, Camps MT, Fernández-Nebro A, González-Escribano MF, Sánchez-Román J, García-Hernández FJ, Castillo MJ, Aguirre MÁ, Gómez-Gracia I, Fernández-Gutiérrez B, Rodríguez-Rodríguez L, Vicente E, Andreu JL, de Castro MF, de la Peña PG, López-Longo FJ, Martínez L, Fonollosa V, Espinosa G, Tolosa C, Pros A, Carballeira MR, Narváez FJ, Rivas MR, Santamaría VO, Díaz B, Trapiella L, del Carmen Freire M, Sousa A, Egurbide MV, Mateo PF, Sáez-Comet L, Díaz F, Hernández V, Beltrán E, Román-Ivorra JA, Grau E, Sancho JJA, García FJB, Oreiro N, Sueiro LF, Zhernakova A, Padyukov L, Alarcón-Riquelme M, Wijmenga C, Brown M, Beretta L, Riemekasten G, Witte T, Hunzelmann N, Kreuter A, Distler JHW, Voskuyl AE, Schuerwegh AJ, Hesselstrand R, Nordin A, Airó P, Lunardi C, Shiels P, van Laar

JM, Herrick A, Worthington J, Denton C, Wigley FM, Hummers LK, Varga J, Hinchcliff ME, Baron M, Hudson M, Pope JE, Furst DE, Khanna D, Phillips K, Schiopu E, Segal BM, Molitor JA, Silver RM, Steen VD, Simms RW, Lafyatis RA, Fessler BJ, Frech TM, AlKassab F, Docherty P, Kaminska E, Khalidi N, Jones HN, Markland J, Robinson D, Broen J, Radstake TRDJ, Fonseca C, Koeleman BP, Martin J (2014) Immunochip Analysis Identifies Multiple Susceptibility Loci for Systemic Sclerosis. The American Journal of Human Genetics 94: 47-61.

377. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, Zhernakova A, Zhernakova DV, Veldink JH, Van den Berg LH, Karjalainen J, Withoff S, Uitterlinden AG, Hofman A, Rivadeneira F, t Hoen PA, Reinmaa E, Fischer K, Nelis M, Milani L, Melzer D, Ferrucci L, Singleton AB, Hernandez DG, Nalls MA, Homuth G, Nauck M, Radke D, Volker U, Perola M, Salomaa V, Brody J, Suchy-Dicey A, Gharib SA, Enquobahrie DA, Lumley T, Montgomery GW, Makino S, Prokisch H, Herder C, Roden M, Grallert H, Meitinger T, Strauch K, Li Y, Jansen RC, Visscher PM, Knight JC, Psaty BM, Ripatti S, Teumer A, Frayling TM, Metspalu A, van Meurs JB, Franke L (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45: 1238-1243.

378. Blankman JL, Simon GM, Cravatt BF (2007) A comprehensive profile of brain enzymes that hydrolyze the endocannabinoid 2-arachidonoylglycerol. Chem Biol 14: 1347-1356.

379. Thomas G, Betters JL, Lord CC, Brown AL, Marshall S, Ferguson D, Sawyer J, Davis MA, Melchior JT, Blume LC, Howlett AC, Ivanova PT, Milne SB, Myers DS, Mrak I, Leber V, Heier C, Taschler U, Blankman JL, Cravatt BF, Lee RG, Crooke RM, Graham MJ, Zimmermann R, Brown HA, Brown JM (2013) The Serine Hydrolase ABHD6 Is a Critical Regulator of the Metabolic Syndrome. CellReports 5: 508-520.

380. Alhouayek M, Masquelier J, Cani PD, Lambert DM, Muccioli GG (2013) Implication of the anti-inflammatory bioactive lipid prostaglandin D2-glycerol ester in the control of macrophage activation and inflammation by ABHD6. Proceedings of the National Academy of Sciences 110: 17558-17563.

381. Tchantchou F, Zhang Y (2013) Selective Inhibition of Alpha/Beta-Hydrolase Domain 6 Attenuates Neurodegeneration, Alleviates Blood Brain Barrier Breakdown, and Improves Functional Recovery in a Mouse Model of Traumatic Brain Injury. Journal of Neurotrauma 30: 565-579.

382. Volk DW, Siegel BI, Verrico CD, Lewis DA (2013) Endocannabinoid metabolism in the prefrontal cortex in schizophrenia. Schizophrenia Research 147: 53-57.

383. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106: 9362-9367.

384. Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, Daly MJ, Neale BM, Sunyaev SR, Lander ES (2014) Searching for missing heritability: designing rare variant association studies. Proc Natl Acad Sci U S A 111: E455-464.

385. Casanova JL, Conley ME, Seligman SJ, Abel L, Notarangelo LD (2014) Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. J Exp Med 211: 2137-2149.

386. Gibson G (2012) Rare and common variants: twenty arguments. Nature Publishing Group 13: 135-145.

387. Nordang GBN, Viken MK, Amundsen SS, Sanchez ES, Flato B, Forre OT, Martin J, Kvien TK, Lie BA (2012) Interferon regulatory factor 5 gene polymorphism confers risk to several rheumatic diseases and correlates with expression of alternative thymic transcripts. Rheumatology 51: 619-626.

388. Sigurdsson S, Göring HHH, Kristjansdottir G, Milani L, Nordmark G, Sandling JK, Eloranta M-L, Feng D, Sangster-Guity N, Gunnarsson I, Svenungsson E, Sturfelt G, Jönsen A, Truedsson L, Barnes BJ, Alm G, Rönnblom L, Syvänen A-C (2008) Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus. Human Molecular Genetics 17: 872-881.

389. Graham DSC, Manku H, Wagner S, Reid J, Timms K, Gutin A, Lanchbury JS, Vyse TJ (2006) Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation. Human Molecular Genetics 16: 579-591.

390. Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, de la Serna AR, Fernandez-Gutierrez B, Castaneda S, Largo R, Joven B, Narvaez J, Navarro F, Marenco JL, Vicario JL, Blanco FJ, Fernandez-Lopez JC, Caliz R, Collado-Escobar MD, Carreno L, Lopez-Longo J, Canete JD, Gomez-Reino JJ, Gonzalez A (2008) Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. Arthritis Rheum 58: 1264-1274.

391. Martin JE, Broen JC, Carmona FD, Teruel M, Simeon CP, Vonk MC, van 't Slot R, Rodriguez-Rodriguez L, Vicente E, Fonollosa V, Ortego-Centeno N, Gonzalez-Gay MA, Garcia-Hernandez FJ, de la Pena PG, Carreira P, Group SS, Voskuyl AE, Schuerwegh AJ, van Riel PLCM, Kreuter A, Witte T, Riemekasten G, Airo P, Scorza R, Lunardi C, Hunzelmann N, Distler JHW, Beretta L, van Laar J, Chee MM, Worthington J, Herrick A, Denton C, Tan FK, Arnett FC, Assassi S, Fonseca C, Mayes MD, Radstake TRDJ, Koeleman BPC, Martin J (2012) Identification of CSK as a systemic sclerosis genetic risk factor through Genome Wide Association Study follow-up. Human Molecular Genetics 21: 2825-2835.

392. Kristjansdottir G, Sandling JK, Bonetti A, Roos IM, Milani L, Wang C, Gustafsdottir SM, Sigurdsson S, Lundmark A, Tienari PJ, Koivisto K, Elovaara I, Pirttila T, Reunanen M, Peltonen L, Saarela J, Hillert J, Olsson T, Landegren U, Alcina A, Fernandez O, Leyva L, Guerrero M, Lucas M, Izquierdo G, Matesanz F, Syvänen A-C (2008) Interferon regulatory factor 5 (IRF5) gene variants are associated with

multiple sclerosis in three distinct populations. Journal of Medical Genetics 45: 362-369.

393. Vosslamber S, van der Voort LF, van den Elskamp IJ, Heijmans R, Aubin C, Uitdehaag BM, Crusius JB, van der Pouw Kraan TC, Comabella M, Montalban X, Hafler DA, De Jager PL, Killestein J, Polman CH, Verweij CL (2011) Interferon regulatory factor 5 gene variants and pharmacological and clinical outcome of Interferonbeta therapy in multiple sclerosis. Genes Immun 12: 466-472.

394. Kottyan LC, Zoller EE, Bene J, Lu X, Kelly JA, Rupert AM, Lessard CJ, Vaughn SE, Marion M, Weirauch MT, Namjou B, Adler A, Rasmussen A, Glenn S, Montgomery CG, Hirschfield GM, Xie G, Coltescu C, Amos C, Li H, Ice JA, Nath SK, Mariette X, Bowman S, Registry UPSasS, Rischmueller M, Lester S, Brun JG, Gøransson LG, Harboe E, Omdal R, Cunninghame Graham DS, Vyse T, Miceli-Richard C, Brennan MT, Lessard JA, Wahren-Herlenius M, Kvarnström M, Illei GG, Witte T, Jonsson R, Eriksson P, Nordmark G, Ng W-F, Registry UPSasS, Anaya J-M, Rhodus NL, Segal BM, Merrill JT, James JA, Guthridge JM, Scofield RH, Alarcón-Riquelme M, Bae S-C, Boackle SA, Criswell LA, Gilkeson G, Kamen DL, Jacob CO, Kimberly R, Brown E, Edberg J, Alarcón GS, Reveille JD, Vila LM, Petri M, Ramsey-Goldman R, Freedman BI, Niewold T, Stevens AM, Tsao BP, Ying J, Mayes MD, Gorlova OY, Wakeland W, Radstake T, Martin E, Martin J, Siminovitch K, Moser Sivils KL, Gaffney PM, Langefeld CD, Harley JB, Kaufman KM (2015) The IRF5-TNPO3 association with systemic lupus

erythematosus has two components that other autoimmune disorders variably share. Human Molecular Genetics 24: 582-596.

395. Hirschhorn JN (2009) Genomewide association studies--illuminating biologic pathways. New England Journal of Medicine 360: 1699-1701.

396. Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, Gupta N, Sklar P, Sullivan PF, Moran JL, Hultman CM, Lichtenstein P, Magnusson P, Lehner T, Shugart YY, Price AL, de Bakker PI, Purcell SM, Sunyaev SR (2012) Exome sequencing and the genetic basis of complex traits. Nat Genet 44: 623-630.

397. Diogo D, Kurreeman F, Stahl EA, Liao KP, Gupta N, Greenberg JD, Rivas MA, Hickey B, Flannick J, Thomson B, Guiducci C, Ripke S, Adzhubey I, Barton A, Kremer JM, Alfredsson L, Consortium of Rheumatology Researchers of North A, Rheumatoid Arthritis Consortium I, Sunyaev S, Martin J, Zhernakova A, Bowes J, Eyre S, Siminovitch KA, Gregersen PK, Worthington J, Klareskog L, Padyukov L, Raychaudhuri S, Plenge RM (2013) Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. Am J Hum Genet 92: 15-27.

398. Min JL, Barrett A, Watts T, Pettersson FH, Lockstone HE, Lindgren CM, Taylor JM, Allen M, Zondervan KT, McCarthy MI (2010) Variability of gene expression profiles in human blood and lymphoblastoid cell lines. BMC Genomics 11: 96.

399. Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, Boucher G, Ripke S, Ellinghaus D, Burtt N, Fennell T, Kirby A, Latiano A, Goyette P, Green T, Halfvarson J, Haritunians T, Korn JM, Kuruvilla F, Lagacé C, Neale B, Lo KS,

Schumm P, Törkvist L, Dubinsky MC, Brant SR, Silverberg MS, Duerr RH, Altshuler D, Gabriel S, Lettre G, Franke A, D'Amato M, McGovern DPB, Cho JH, Rioux JD, Xavier RJ, Daly MJ (2011) Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nature Genetics 43: 1066-1073.

400. Guthridge JM, Lu R, Sun H, Sun C, Wiley GB, Dominguez N, Macwana SR, Lessard CJ, Kim-Howard X, Cobb BL, Kaufman KM, Kelly JA, Langefeld CD, Adler AJ, Harley IT, Merrill JT, Gilkeson GS, Kamen DL, Niewold TB, Brown EE, Edberg JC, Petri MA, Ramsey-Goldman R, Reveille JD, Vila LM, Kimberly RP, Freedman BI, Stevens AM, Boackle SA, Criswell LA, Vyse TJ, Behrens TW, Jacob CO, Alarcon-Riquelme ME, Sivils KL, Choi J, Joo YB, Bang SY, Lee HS, Bae SC, Shen N, Qian X, Tsao BP, Scofield RH, Harley JB, Webb CF, Wakeland EK, James JA, Nath SK, Graham RR, Gaffney PM (2014) Two functional lupus-associated BLK promoter variants control cell-type- and developmental-stage-specific transcription. Am J Hum Genet 94: 586-598.

401. Pasaniuc B, Rohland N, McLaren PJ, Garimella K, Zaitlen N, Li H, Gupta N, Neale BM, Daly MJ, Sklar P, Sullivan PF, Bergen S, Moran JL, Hultman CM, Lichtenstein P, Magnusson P, Purcell SM, Haas DW, Liang L, Sunyaev S, Patterson N, de Bakker PI, Reich D, Price AL (2012) Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. Nat Genet 44: 631-635.

402. Cusick ME, Yu H, Smolyar A, Venkatesan K, Carvunis AR, Simonis N, Rual JF, Borick H, Braun P, Dreze M, Vandenhaute J, Galli M, Yazaki J, Hill DE, Ecker JR, Roth FP, Vidal M (2009) Literature-curated protein interaction datasets. Nat Methods 6: 39-46.

403. Kichaev G, Yang W-Y, Lindstrom S, Hormozdiari F, Eskin E, Price AL, Kraft P, Pasaniuc B (2014) Integrating Functional Data to Prioritize Causal Variants in Statistical Fine-Mapping Studies. PLoS Genetics 10: e1004722.

404. Grossman SR, Andersen KG, Shlyakhter I, Tabrizi S, Winnicki S, Yen A, Park DJ, Griesemer D, Karlsson EK, Wong SH, Cabili M, Adegbola RA, Bamezai RNK, Hill AVS, Vannberg FO, Rinn JL, Lander ES, Schaffner SF, Sabeti PC, Project G (2013) Identifying Recent Adaptations in Large-Scale Genomic Data. Cell 152: 703-713.

405. Goldstein DB, Allen A, Keebler J, Margulies EH, Petrou S, Petrovski S, Sunyaev S (2013) Sequencing studies in human genetics: design and interpretation. Nature Reviews Genetics 14: 460-470.

406. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57-74.

407. Hardison RC (2012) Genome-wide Epigenetic Data Facilitate Understanding of Disease Susceptibility Association Studies. Journal of Biological Chemistry 287: 30932-30940.

408. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M (2012) Linking disease associations with regulatory information in the human genome. Genome Research 22: 1748-1759.

409. Akhtar-Zaidi B, Cowper-Sal-lari R, Corradin O, Saiakhova A, Bartels CF, Balasubramanian D, Myeroff L, Lutterbaugh J, Jarrar A, Kalady MF, Willis J, Moore JH, Tesar PJ, Laframboise T, Markowitz S, Lupien M, Scacheri PC (2012) Epigenomic enhancer profiling defines a signature of colon cancer. Science 336: 736-739.

410. Chen L, Kostadima M, Martens JH, Canu G, Garcia SP, Turro E, Downes K, Macaulay IC, Bielczyk-Maczynska E, Coe S, Farrow S, Poudel P, Burden F, Jansen SB, Astle WJ, Attwood A, Bariana T, de Bono B, Breschi A, Chambers JC, Consortium B, Choudry FA, Clarke L, Coupland P, van der Ent M, Erber WN, Jansen JH, Favier R, Fenech ME, Foad N, Freson K, van Geet C, Gomez K, Guigo R, Hampshire D, Kelly AM, Kerstens HH, Kooner JS, Laffan M, Lentaigne C, Labalette C, Martin T, Meacham S, Mumford A, Nurnberg S, Palumbo E, van der Reijden BA, Richardson D, Sammut SJ, Slodkowicz G, Tamuri AU, Vasquez L, Voss K, Watt S, Westbury S, Flicek P, Loos R, Goldman N, Bertone P, Read RJ, Richardson S, Cvejic A, Soranzo N, Ouwehand WH, Stunnenberg HG, Frontini M, Rendon A (2014) Transcriptional diversity during lineage commitment of human blood progenitors. Science 345: 1251033.

411. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456: 470-476.

412. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature Genetics 40: 1413-1415.

413. Corominas R, Yang X, Lin GN, Kang S, Shen Y, Ghamsari L, Broly M, Rodriguez M, Tam S, Trigg SA, Fan C, Yi S, Tasan M, Lemmens I, Kuang X, Zhao N, Malhotra D, Michaelson JJ, Vacic V, Calderwood MA, Roth FP, Tavernier J, Horvath S, Salehi-Ashtiani K, Korkin D, Sebat J, Hill DE, Hao T, Vidal M, Iakoucheva LM (2014) Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. Nature Communications 5: 1-12.

414. Shi L, Zhang Z, Yu AM, Wang W, Wei Z, Akhter E, Maurer K, Costa Reis P, Song L, Petri M, Sullivan KE (2014) The SLE transcriptome exhibits evidence of chronic endotoxin exposure and has widespread dysregulation of non-coding and coding RNAs. PLoS One 9: e93846.

415. Stone RC, Du P, Feng D, Dhawan K, Rönnblom L, Eloranta M-L, Donnelly R, Barnes BJ (2013) RNA-Seq for Enrichment and Analysis of IRF5 Transcript Expression in SLE. PLoS ONE 8: e54487.

416. Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology 32: 347-355.

417. Gupta RM, Musunuru K (2014) Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. Journal of Clinical Investigation 124: 4154-4161.

418. Chu HW, Rios C, Huang C, Wesolowska-Andersen A, Burchard EG, O'Connor BP, Fingerlin TE, Nichols D, Reynolds SD, Seibold MA (2015) CRISPR-Cas9 mediated gene knockout in primary human airway epithelial cells reveals a pro-inflammatory role for MUC18. Gene Ther.

419. Yao L, Tak YG, Berman BP, Farnham PJ (2014) Functional annotation of colon cancer risk SNPs. Nature Communications 5: 1-13.

420. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell 153: 910-918.

421. Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN (2014) Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science 345: 1184-1188.

422. Huang X, Wang Y, Yan W, Smith C, Ye Z, Wang J, Gao Y, Mendelsohn L, Cheng L (2015) Production of Gene-Corrected Adult Beta Globin Protein in Human Erythrocytes Differentiated from Patient iPSCs After Genome Editing of the Sickle Point Mutation. Stem Cells 33: 1470-1479.

423. Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, Kang Y, Zhao X, Si W, Li W, Xiang AP, Zhou J, Guo X, Bi Y, Si C, Hu B, Dong G, Wang H, Zhou Z, Li T, Tan T, Pu X, Wang F, Ji S, Zhou Q, Huang X, Ji W, Sha J (2014) Generation of

Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. Cell 156: 836-843.

424. Kulis M, Merkel A, Heath S, s ACQo, Schuyler RP, Castellano G, Beekman Ree, Raineri E, Esteve A, Clot G, Verdaguer-Dot Ner, Duran-Ferrer Mi, ol NRn, Vilarrasa-Blasi R, Ecker S, Pancaldi V, Rico D, Agueda L, Blanc J, Richardson D, Clarke L, Datta A, Pascual M, Agirre X, Prosper F, Alignani D, Paiva B, Caron G, Fest T, Muench MO, Fomin ME, Lee S-T, Wiemels JL, Valencia A, Gut M, Flicek P, Stunnenberg HG, Siebert R, ppers RKu, Gut IG, Campo Eia, n-Subero JeIMi (2015) Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nature Genetics: 1-14.

425. Auwerx J, Avner P, Baldock R, Ballabio A, Balling R, Barbacid M, Berns A, Bradley A, Brown S, Carmeliet P, Chambon P, Cox R, Davidson D, Davies K, Duboule D, Forejt J, Granucci F, Hastie N, de Angelis MH, Jackson I, Kioussis D, Kollias G, Lathrop M, Lendahl U, Malumbres M, von Melchner H, Muller W, Partanen J, Ricciardi-Castagnoli P, Rigby P, Rosen B, Rosenthal N, Skarnes B, Stewart AF, Thornton J, Tocchini-Valentini G, Wagner E, Wahli W, Wurst W (2004) The European dimension for the mouse genome mutagenesis program. Nat Genet 36: 925-927.

426. Rickert RC, Roes J, Rajewsky K (1997) B lymphocyte-specific, Cre-mediated mutagenesis in mice. Nucleic Acids Research 25: 1317-1318.

427. Satoh M, Reeves WH (1994) Induction of lupus-associated autoantibodies in BALB/c mice by intraperitoneal injection of pristane. The Journal of experimental medicine 180: 2341-2346.

428. Satoh M, Kumar A, Kanwar YS, Reeves WH (1995) Anti-nuclear antibody production and immune-complex glomerulonephritis in BALB/c mice treated with pristane. Proceedings of the National Academy of Sciences 92: 10934-10938.

429. Leiss H, Niederreiter B, Bandur T, Schwarzecker B, Bluml S, Steiner G, Ulrich W, Smolen J, Stummvoll G (2013) Pristane-induced lupus as a model of human lupus arthritis: evolvement of autoantibodies, internal organ and joint inflammation. Lupus 22: 778-792.

430. Reeves WH, Lee PY, Weinstein JS, Satoh M, Lu L (2009) Induction of autoimmunity by pristane and other naturally occurring hydrocarbons. Trends in Immunology 30: 455-464.

431. Deng Y, Tsao BP (2014) Advances in lupus genetics and epigenetics. Curr Opin Rheumatol 26: 482-492.

432. Zhang Y, Zhao M, Sawalha AH, Richardson B, Lu Q (2013) Impaired DNA methylation and its mechanisms in CD4(+)T cells of systemic lupus erythematosus. J Autoimmun 41: 92-99.

433. Absher DM, Li X, Waite LL, Gibson A, Roberts K, Edberg J, Chatham WW, Kimberly RP (2013) Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. PLoS Genet 9: e1003678.

434. Dozmorov MG, Wren JD, Alarcon-Riquelme ME (2014) Epigenomic elements enriched in the promoters of autoimmunity susceptibility genes. Epigenetics 9: 276-285.

435. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S, Graham RR, Manoharan A, Ortmann W, Bhangale T, Denny JC, Carroll RJ, Eyler AE, Greenberg JD, Kremer JM, Pappas DA, Jiang L, Yin J, Ye L, Su DF, Yang J, Xie G, Keystone E, Westra HJ, Esko T, Metspalu A, Zhou X, Gupta N, Mirel D, Stahl EA, Diogo D, Cui J, Liao K, Guo MH, Myouzen K, Kawaguchi T, Coenen MJ, van Riel PL, van de Laar MA, Guchelaar HJ, Huizinga TW, Dieude P, Mariette X, Bridges SL, Jr., Zhernakova A, Toes RE, Tak PP, Miceli-Richard C, Bang SY, Lee HS, Martin J, Gonzalez-Gay MA, Rodriguez-Rodriguez L, Rantapaa-Dahlqvist S, Arlestig L, Choi HK, Kamatani Y, Galan P, Lathrop M, Eyre S, Bowes J, Barton A, de Vries N, Moreland LW, Criswell LA, Karlson EW, Taniguchi A, Yamada R, Kubo M, Liu JS, Bae SC, Worthington J, Padyukov L, Klareskog L, Gregersen PK, Raychaudhuri S, Stranger BE, De Jager PL, Franke L, Visscher PM, Brown MA, Yamanaka H, Mimori T, Takahashi A, Xu H, Behrens TW, Siminovitch KA, Momohara S, Matsuda F, Yamamoto K, Plenge RM (2014) Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506: 376-381.

436. Enyedy EJ, Mitchell JP, Nambiar MP, Tsokos GC (2001) Defective FcgammaRIIb1 signaling contributes to enhanced calcium response in B cells from patients with systemic lupus erythematosus. Clin Immunol 101: 130-135.

437. Clark MR, Massenburg D, Siemasko K, Hou P, Zhang M (2004) B-cell antigen receptor signaling requirements for targeting antigen to the MHC class II presentation pathway. Current Opinion in Immunology 16: 382-387.

438. O'Neill SK, Veselits ML, Zhang M, Labno C, Cao Y, Finnegan A, Uccellini M, Alegre M-L, Cambier JC, Clark MR (2009) Endocytic sequestration of the B cell antigen receptor and toll-like receptor 9 in anergic cells. Proceedings of the National Academy of Sciences of the United States of America 106: 6262-6267.

439. Webb CF, Das C, Coffman RL, Tucker PW (1989) Induction of immunoglobulin mu mRNA in a B cell transfectant stimulated with interleukin-5 and a T-dependent antigen. Journal of immunology (Baltimore, Md : 1950) 143: 3934-3939.

440. Ferrari G, Knight AM, Watts C, Pieters J (1997) Distinct intracellular compartments involved in invariant chain degradation and antigenic peptide loading of major histocompatibility complex (MHC) class II molecules. The Journal of Cell Biology 139: 1433-1446.

441. Courtney AH, Bennett NR, Zwick DB, Hudon J, Kiessling LL (2014) Synthetic antigens reveal dynamics of BCR endocytosis during inhibitory signaling. ACS Chem Biol 9: 202-210.

442. Le Roux D, Lankar D, Yuseff M-I, Vascotto F, Yokozeki T, Faure-André G, Mougneau E, Glaichenhaus N, Manoury B, Bonnerot C, Lennon-Duménil A-M (2007) Syk-dependent actin dynamics regulate endocytic trafficking and processing of antigens internalized through the B-cell receptor. Molecular Biology of the Cell 18: 3451-3462.

443. Katkere B, Rosa S, Drake JR (2012) The Syk-binding ubiquitin ligase c-Cbl mediates signaling-dependent B cell receptor ubiquitination and B cell receptor-mediated antigen processing and presentation. J Biol Chem 287: 16636-16644.

444. Clark MR, Johnson SA, Cambier JC (1994) Analysis of Ig-alpha-tyrosine kinase interaction reveals two levels of binding specificity and tyrosine phosphorylated Ig-alpha stimulation of Fyn activity. The EMBO Journal.

445. Niiro H, Clark EA (2002) Decision making in the Immune System: Regulation of B-cell fate by antigen-receptor signals. Nature Reviews Immunology 2: 945-956.

446. Schmitz R, Baumann G, Gram H (1996) Catalytic specificity of phosphotyrosine kinases Blk, Lyn, c-Src and Syk as assessed by phage display. J Mol Biol 260: 664-677.

447. Jang C, Machtaler S, Matsuuchi L (2010) The role of Ig-alpha/beta in B cell antigen receptor internalization. Immunol Lett 134: 75-82.

448. Busman-Sahay K, Drake L, Sitaram A, Marks M, Drake JR (2013) Cis and trans regulatory mechanisms control AP2-mediated B cell receptor endocytosis via select tyrosine-based motifs. PLoS One 8: e54938.

449. Raiborg C, Stenmark H (2009) The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. Nature 458: 445-452.

450. Slagsvold T, Aasland R, Hirano S, Bache KG, Raiborg C, Trambaiolo D, Wakatsuki S, Stenmark H (2005) Eap45 in mammalian ESCRT-II binds ubiquitin via a phosphoinositide-interacting GLUE domain. J Biol Chem 280: 19600-19606.

451. Veselits M, Tanaka A, Lipkowitz S, O'Neill S, Sciammas R, Finnegan A, Zhang J, Clark MR (2014) Recruitment of Cbl-b to B cell antigen receptor couples antigen recognition to Toll-like receptor 9 activation in late endosomes. PLoS One 9: e89792.

452. Greenwald RJ, Freeman GJ, Sharpe AH (2005) The B7 family revisited. Annu Rev Immunol 23: 515-548.

453. Krieg C, Boyman O, Fu YX, Kaye J (2007) B and T lymphocyte attenuator regulates CD8+ T cell-intrinsic homeostasis and memory cell generation. Nat Immunol 8: 162-171.

454. Siemasko K, Eisfelder BJ, Williamson E, Kabak S, Clark MR (1998) Cutting edge: signals from the B lymphocyte antigen receptor regulate MHC class II containing late endosomes. Journal of immunology (Baltimore, Md : 1950) 160: 5203-5208.

455. Nemazee D, Buerki K (1989) Clonal deletion of autoreactive B lymphocytes in bone marrow chimeras. Proceedings of the National Academy of Sciences 86: 8039-8043.

456. Erikson J, Radic MZ, Camper SA, Hardy RR, Carmack C, Weigert M (1991) Expression of anti-DNA immunoglobulin transgenes in non-autoimmune mice. Nature 349: 331-334.

457. Goodnow CC, Vinuesa CG, Randall KL, Mackay F, Brink R (2010) Control systems and decision making for antibody production. Nat Immunol 11: 681-688.

458. Yurasov S (2005) Defective B cell tolerance checkpoints in systemic lupus erythematosus. Journal of Experimental Medicine 201: 703-711.

459. Merrell KT, Benschop RJ, Gauld SB, Aviszus K, Decote-Ricardo D, Wysocki LJ, Cambier JC (2006) Identification of anergic B cells within a wild-type repertoire. Immunity 25: 953-962.

460. Duty JA, Szodoray P, Zheng NY, Koelsch KA, Zhang Q, Swiatkowski M, Mathias M, Garman L, Helms C, Nakken B, Smith K, Farris AD, Wilson PC (2009) Functional anergy in a subpopulation of naive B cells from healthy humans that express autoreactive immunoglobulin receptors. J Exp Med 206: 139-151.

461. Isnardi I, Ng YS, Menard L, Meyers G, Saadoun D, Srdanovic I, Samuels J, Berman J, Buckner JH, Cunningham-Rundles C, Meffre E (2010) Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones. Blood 115: 5026-5036.

462. Yarkoni Y, Getahun A, Cambier JC (2010) Molecular underpinning of B-cell anergy. Immunological reviews 237: 249-263.

463. Zikherman J, Parameswaran R, Weiss A (2012) Endogenous antigen tunes the responsiveness of naive B cells but not T cells. Nature 489: 160-164.

464. Vaughn SE, Foley C, Lu X, Patel ZH, Zoller EE (2014) Lupus risk variants in the PXK locus alter B-cell receptor internalization. Frontiers in ….

465. Lu X, Zoller EE, Weirauch MT, Wu Z, Namjou B, Williams AH, Ziegler JT, Comeau ME, Marion MC, Glenn SB, Adler A, Shen N, Nath SK, Stevens AM, Freedman BI, Tsao BP, Jacob CO, Kamen DL, Brown EE, Gilkeson GS, Alarcon GS, Reveille JD, Anaya JM, James JA, Sivils KL, Criswell LA, Vila LM, Alarcon-Riquelme ME, Petri M, Scofield RH, Kimberly RP, Ramsey-Goldman R, Joo YB,

Choi J, Bae SC, Boackle SA, Graham DC, Vyse TJ, Guthridge JM, Gaffney PM, Langefeld CD, Kelly JA, Greis KD, Kaufman KM, Harley JB, Kottyan LC (2015) Lupus Risk Variant Increases pSTAT1 Binding and Decreases ETS1 Expression. Am J Hum Genet 96: 731-739.